



Perbandingan *Random Forest Regressor* Dan *Decision Tree Regressor* Untuk Prediksi Hasil Panen

Rizki Faizal¹, Asrul Abdullah*², Menur Wahyu Pangestika³

Email: ¹211220007@unmuhpnk.ac.id, ²asrul.abdullah@unmuhpnk.ac.id, ³menur.wahyu@unmuhpnk.ac.id

¹Program Studi Teknik Informatika, Fakultas Teknik dan Ilmu Komputer, Universitas Muhammadiyah Pontianak

²Program Studi Teknik Informatika, Fakultas Teknik dan Ilmu Komputer, Universitas Muhammadiyah Pontianak

³Program Studi Teknik Informatika, Fakultas Teknik dan Ilmu Komputer, Universitas Muhammadiyah Pontianak

Diterima: 6 Agustus 2025 | Direvisi: - | Disetujui: 5 September 2025

©2020 Program Studi Teknik Informatika Fakultas Ilmu Komputer,
Universitas Muhammadiyah Riau, Indonesia

Abstrak

Ketidakpastian hasil panen akibat faktor lingkungan menjadi tantangan besar dalam sektor pertanian Indonesia. Penelitian ini bertujuan membandingkan kinerja algoritma *Random Forest Regressor* dan *Decision Tree Regressor* dalam memprediksi hasil panen tanaman budidaya. Dataset yang digunakan berasal dari dataset terbuka sebanyak 300.000 baris data dengan fitur seperti jenis tanaman, jenis tanah, curah hujan, penggunaan pupuk, irigasi, dan kondisi cuaca. Sistem dikembangkan menggunakan Python dan Streamlit. Metodologi mencakup prapemrosesan data, pelatihan model, serta evaluasi menggunakan metrik *Mean Absolute Error (MAE)*. Hasil pengujian menunjukkan bahwa *Decision Tree Regressor* memiliki *MAE* lebih rendah (0,43) dibandingkan *Random Forest Regressor* (0,48), sehingga menghasilkan prediksi yang lebih akurat pada data ini. Analisis fitur menunjukkan bahwa curah hujan dan jenis tanaman merupakan faktor paling berpengaruh. Meskipun *Random Forest* dikenal lebih stabil dalam banyak kasus, penelitian ini menunjukkan bahwa *Decision Tree* dapat lebih unggul dalam konteks dataset yang digunakan. Sistem yang dibangun diharapkan dapat membantu petani dan pemangku kebijakan dalam merencanakan produksi pertanian secara lebih efisien dan berbasis data.

Kata kunci: *Decision Tree, Hasil Panen, MAE, Prediksi, Random Forest*

Comparison Of *Random Forest Regressor* And *Decision Tree Regressor* For Crop Yield Prediction

Abstract

Uncertainty in crop yields due to environmental factors remains a major challenge in Indonesia's agricultural sector. This study aims to compare the performance of the *Random Forest Regressor* and *Decision Tree Regressor* algorithms in predicting cultivated crop yields. The dataset used was sourced from Kaggle, consisting of 300,000 rows with features such as crop type, soil type, rainfall, fertilizer use, irrigation, and weather conditions. The system was developed using Python and Streamlit. The methodology includes data preprocessing, model training, and evaluation using the *Mean Absolute Error (MAE)* metric. The test results show that the *Decision Tree Regressor* achieved a lower *MAE* (0.43) compared to the *Random Forest Regressor* (0.48), resulting in more accurate predictions on this dataset. Feature analysis indicates that rainfall and crop type are the most influential factors. Although *Random Forest* is generally known for its stability, this study demonstrates that *Decision Tree* can outperform it within the context of the dataset used. The developed system is expected to assist farmers and policymakers in planning agricultural production more efficiently and in a data-driven manner.

Keywords: *Decision Tree, MAE, Prediction, Random Forest, Yield*

1. PENDAHULUAN

Berdasarkan data dari Badan Pusat Statistik (BPS), padi, jagung, dan kedelai merupakan komoditas utama yang berkontribusi besar terhadap ketahanan budidaya nasional[1]. Tanaman yang dikaji dalam penelitian ini mencakup padi, kapas, kedelai, jagung, gandum, dan jelai. Padi, jagung, dan kedelai merupakan komoditas utama yang banyak dibudidayakan di Indonesia, sedangkan gandum dan jelai merupakan tanaman yang umumnya tumbuh di wilayah beriklim sedang dan belum dibudidayakan secara luas di Indonesia[2]. Sementara itu, kapas juga dibudidayakan meskipun dalam skala terbatas dan memiliki potensi untuk dikembangkan di beberapa wilayah[3]. Sektor pertanian berperan strategis dalam perekonomian Indonesia sebagai penyedia pangan, bahan baku, tenaga kerja, dan devisa. Ketidakpastian hasil panen menjadi tantangan utama yang berdampak pada ketahanan pangan dan stabilitas ekonomi. Oleh karena itu, prediksi produksi tanaman budidaya diperlukan untuk mendukung perencanaan pertanian yang efisien dan berkelanjutan[4].

Berbagai penelitian sebelumnya telah membuktikan efektivitas algoritma pembelajaran mesin dalam memprediksi hasil panen. Penelitian oleh Swain et al. menggunakan algoritma *Random Forest Regressor* dan mencatat nilai R^2 sebesar 0,98 dan MAE 0,1088, menunjukkan bahwa model mampu menjelaskan 98% variasi hasil panen berdasarkan faktor lingkungan dan input[5]. Penelitian lokal oleh Sulfayanti dan Wildayani juga menunjukkan hasil serupa, dengan akurasi 95,11%, $R^2 = 0,99$, serta $MAPE = 4,884\%$ dan $RMSE = 0,250$, membuktikan keandalan *Random Forest* dalam konteks pertanian Indonesia[6]. Selain penelitian di bidang pertanian, sejumlah studi di bidang lain juga menunjukkan keandalan algoritma *Random Forest* dalam menangani permasalahan prediksi berbasis data. Penelitian oleh Mualfah et al. menunjukkan bahwa penerapan algoritma *Random Forest* dengan teknik *SMOTE* dalam mendeteksi penyakit stroke menghasilkan *precision*, *recall*, dan *f1-score* masing-masing sebesar 0,92, 0,91, dan 0,91, meskipun akurasi menurun dari 0,98 menjadi 0,91 setelah penyeimbangan data[7]. Sementara itu, penelitian oleh Handayani dan Medikawati menegaskan bahwa kombinasi *Random Forest* dan *SMOTE* mampu mencapai akurasi, *precision*, *recall*, dan *f1-score* sebesar 95%, 95%, 100%, dan 97% dalam klasifikasi penyakit stroke otak, mengungguli algoritma lain seperti *KNN* dan *Logistic Regression* dalam kondisi data tidak seimbang[8]. Di sisi lain, algoritma *Decision Tree Regressor* juga menunjukkan potensi prediktif yang baik. Penelitian oleh Irawan dalam prediksi kualitas dan produksi padi di Pulau Jawa menunjukkan nilai awal $R^2 = 1$ (sebelum pembersihan data), meskipun model mengalami *overfitting*. Setelah data dibersihkan dari *outlier*, performa model meningkat dengan R^2 dan $EVS = 0,89$ serta $MSE = 280.703.454,88$ [9].

Tidak hanya pada sektor pertanian, potensi algoritma ini juga terlihat dalam konteks yang berbeda. Penelitian oleh Ramadhan et al. menerapkan *Decision Tree Regression* untuk memprediksi harga rumah di Provinsi Jawa Barat. Hasilnya menunjukkan bahwa model ini menghasilkan nilai R^2 sebesar 0,86, lebih tinggi dibandingkan regresi linier berganda dan *Support Vector Regression*, dengan *Mean Absolute Error (MAE)* sebesar 91.436.523, yang menunjukkan performa prediksi yang akurat dan dapat diandalkan dalam sektor properti[10]. Sementara itu, studi oleh Xiang Huang menunjukkan bahwa *Decision Tree* fleksibel untuk klasifikasi dan regresi, serta efektif dalam menilai risiko kredit berdasarkan histori dan durasi pinjaman di sektor keuangan[11]. Hal ini menegaskan bahwa meskipun sensitif terhadap *outlier*, algoritma *Decision Tree* masih relevan digunakan.

Berdasarkan hal tersebut, penelitian ini bertujuan untuk membangun sistem prediksi hasil panen dengan membandingkan dua algoritma supervised learning, yaitu *Random Forest Regressor* dan *Decision Tree Regressor*. Berbeda dari penelitian sebelumnya yang umumnya hanya menggunakan satu algoritma atau terbatas pada fitur dasar, penelitian ini menggunakan pendekatan komparatif dengan tambahan fitur seperti *Soil_Type*, *Irrigation_Used*, dan *Weather_Condition*. Tujuan utama dari penelitian ini adalah untuk mengetahui algoritma mana yang lebih akurat dan andal dalam memprediksi hasil panen berdasarkan faktor lingkungan dan kondisi pertanian secara kompleks dan realistis.

2. METODE PENELITIAN

Metodologi yang digunakan dalam penelitian ini disusun secara sistematis untuk mencapai tujuan yang telah ditetapkan. Tahapan penelitian meliputi proses identifikasi masalah hingga penarikan kesimpulan, sebagaimana ditampilkan pada Gambar 3.1. Setiap langkah dalam alur tersebut dirancang untuk memastikan bahwa proses prediksi hasil panen dapat dilakukan secara terstruktur, mulai dari pengumpulan data, perancangan sistem, implementasi algoritma, hingga pengujian model.



Gambar 2.1 Diagram Alur Metodologi Penelitian

2.1. Identifikasi Masalah

Dalam penelitian ini, penulis menggunakan pendekatan *machine learning* untuk memprediksi hasil panen dengan menerapkan algoritma *Random Forest Regressor* dan *Decision Tree Regressor*. Proses ini mencakup analisis data dan penerapan model prediktif guna mengidentifikasi pola yang berkaitan dengan faktor lingkungan dan produktivitas pertanian. Tujuan utama dari metode ini adalah untuk membandingkan efektivitas kedua algoritma dalam mengatasi ketidakseimbangan data, sensitivitas terhadap parameter, efisiensi komputasi, serta tingkat akurasi prediksi pada berbagai jenis tanaman.

2.2. Studi Literatur

Kajian literatur dilakukan untuk mengidentifikasi dan menganalisis penelitian terdahulu yang relevan dengan penerapan algoritma *Decision Tree Regressor* dan *Random Forest Regressor* dalam konteks prediksi hasil panen. Literatur yang digunakan mencakup jurnal ilmiah dan sumber data pendukung yang menjadi dasar dalam perancangan eksperimen, pemilihan metode, serta evaluasi performa model.

2.3. Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari situs *Kaggle*, dengan tautan: <https://www.kaggle.com/datasets/samuelotiattakorah/agriculture-crop-yield/data>. Dataset ini berisi sekitar 1.000.000 sampel data pertanian dan bertujuan untuk memprediksi hasil panen (dalam ton per hektar). Peneliti mengambil hanya 300.000 data sampel berdasarkan beberapa faktor seperti jenis tanaman, curah hujan, dan kondisi cuaca.

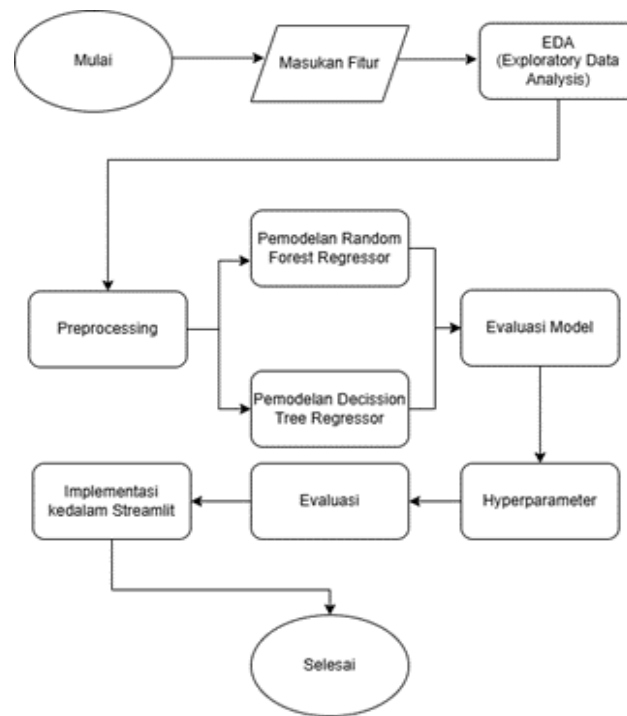
Tabel 2.1 Fitur dan Keterangan Dataset

No	Fitur	Keterangan	Jenis Data
1	<i>Soil_Type</i>	Jenis tanah	Kategorikal
2	<i>Crop</i>	Jenis tanaman	Kategorikal
3	<i>Rainfall_mm</i>	Curah hujan (dalam Milimeter)	Numerikal
4	<i>Fertilizer_Used</i>	Menunjukkan apakah pupuk digunakan atau tidak	Boolean
5	<i>Irrigation_Used</i>	Menunjukkan apakah sistem irigasi digunakan selama masa tanam	Boolean
6	<i>Weather_Condition</i>	Kondisi cuaca utama selama musim tanam	Kategorikal
7	<i>Yield_tons_per_hectare</i>	Hasil panen total yang dihasilkan	Numerikal

2.4. Perancangan

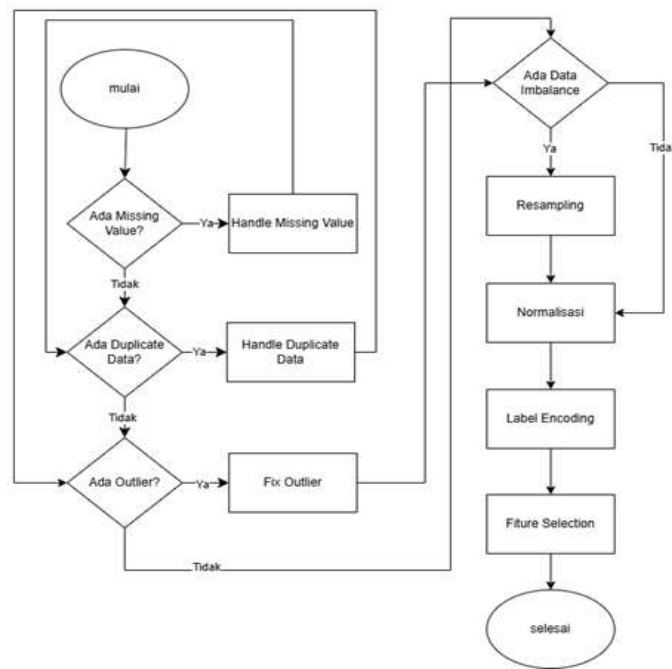
Tahapan perancangan dalam penelitian ini bertujuan untuk membangun sistem prediksi hasil panen berdasarkan algoritma *machine learning*. Tahapan yang pertama adalah perancangan diagram alur proses, yang kedua perancangan proses preprocessing, dan yang ketiga perancangan proses pembangunan model.

Pertama perancangan alur proses sistem prediksi hasil panen dilakukan secara terstruktur untuk memastikan setiap tahapan berjalan efisien dan sesuai tujuan. Diagram alur pada Gambar 2.2 menggambarkan tahapan utama dalam penelitian ini, dimulai dari proses awal yaitu input fitur dan analisis data eksploratif (*EDA*). Selanjutnya dilakukan prapemrosesan data sebelum masuk ke tahap pemodelan menggunakan dua algoritma, yaitu *Random Forest Regressor* dan *Decision Tree Regressor*. Hasil dari kedua model kemudian dievaluasi berdasarkan metrik performa, serta dilakukan penyesuaian hyperparameter untuk mengoptimalkan hasil prediksi. Tahap akhir dari proses ini adalah implementasi ke dalam platform antarmuka web menggunakan *Streamlit*, hingga sistem siap digunakan.



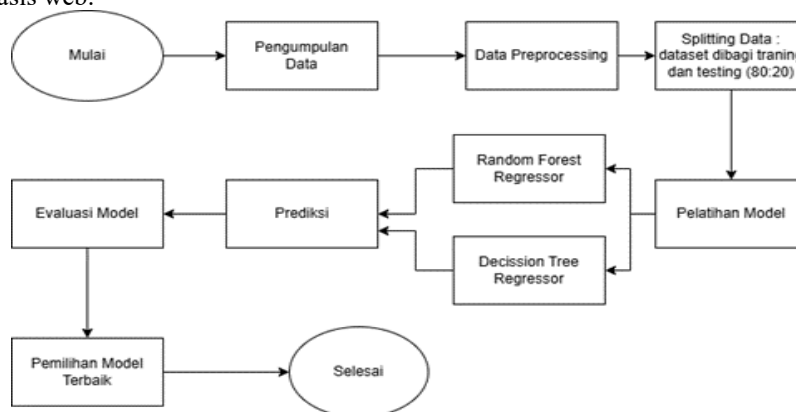
Gambar 2.2 Diagram Alur Proses Sistem Prediksi Hasil Panen

Kedua tahap preprocessing merupakan proses penting dalam menyiapkan data agar model *machine learning* dapat belajar secara optimal. Pada penelitian ini, preprocessing mencakup beberapa langkah, seperti ditunjukkan pada Gambar 2.3. Langkah pertama adalah pengecekan dan penanganan *missing value*, data duplikat, dan *outlier*. Data yang tidak lengkap atau mengandung informasi ganda dapat memengaruhi akurasi model, sehingga perlu disaring terlebih dahulu. Setelah itu, dilakukan pengecekan apakah data memiliki distribusi kelas yang tidak seimbang (*data imbalance*). Jika ya, maka diterapkan teknik *resampling* untuk menyeimbangkan distribusi kelas. Langkah selanjutnya adalah proses normalisasi pada fitur numerik agar berada dalam skala yang seragam, kemudian dilakukan *label encoding* pada fitur kategorikal seperti *Soil_Type*, *Crop*, dan *Weather_Condition*. Terakhir, dilakukan seleksi fitur (*feature selection*) untuk mengidentifikasi fitur-fitur yang paling relevan terhadap target prediksi.



Gambar 2.3 Diagram Proses Preprocessing Data

Ketiga proses pembangunan model dimulai setelah data dikumpulkan dan melalui tahap preprocessing. Dataset kemudian dibagi menjadi dua bagian, yaitu data latih dan data uji dengan rasio 80:20 menggunakan teknik *train-test split*. Selanjutnya, model dibangun dengan menggunakan dua algoritma supervised learning, yaitu *Random Forest Regressor* dan *Decision Tree Regressor*, seperti ditunjukkan pada Gambar 2.4. Setelah proses pelatihan, kedua model digunakan untuk melakukan prediksi hasil panen berdasarkan fitur yang tersedia. Keluaran dari prediksi kemudian dievaluasi menggunakan metrik *Mean Absolute Error (MAE)* untuk mengukur tingkat akurasi. Berdasarkan hasil evaluasi tersebut, dilakukan pemilihan model terbaik, yaitu model dengan nilai *MAE* paling rendah. Model terpilih selanjutnya digunakan sebagai dasar dalam pengembangan sistem prediksi hasil panen berbasis web.



Gambar 2.4 Diagram Proses Pembangunan Model Prediksi Hasil Panen

2.5. Implementasi

Penelitian ini menggunakan algoritma *Random Forest Regressor* dan *Decision Tree Regressor* untuk memprediksi hasil panen. Setelah preprocessing dan pembagian data, model dievaluasi menggunakan metrik regresi: *MAE*, *MSE*, dan R^2 , dengan *MAE* sebagai metrik utama evaluasi. Menurut Dumre et al., *MAE* dipilih karena lebih stabil terhadap outlier dan memberikan hasil evaluasi dalam satuan yang sama dengan nilai target, sehingga mudah diinterpretasikan, terutama pada data pertanian yang cenderung seimbang[12].

2.6. Pengujian

Pengujian dilakukan untuk menjawab rumusan masalah, yaitu membandingkan efektivitas algoritma *Random Forest Regressor* dan *Decision Tree Regressor* dalam membangun sistem prediksi hasil panen yang akurat dan efisien. Model yang telah dilatih diuji menggunakan data uji dan dievaluasi berdasarkan metrik *Mean Absolute Error (MAE)*. Pengujian difokuskan pada dua hal:

- a) Menilai kemampuan masing-masing algoritma dalam menghasilkan prediksi yang akurat dan stabil.
 - b) Membandingkan nilai *MAE* guna menentukan algoritma dengan tingkat kesalahan prediksi terendah.
- Hasil evaluasi digunakan sebagai dasar pemilihan model terbaik untuk sistem prediksi hasil panen berbasis data.

2.7. Kesimpulan

Hasil penelitian menunjukkan bahwa salah satu model mampu memberikan performa yang lebih baik dalam memprediksi hasil panen. Hal ini ditunjukkan oleh nilai evaluasi yang lebih optimal, yang mencerminkan tingkat akurasi dan stabilitas yang lebih baik dalam mengolah data. Dengan demikian, pendekatan ini berpotensi digunakan untuk meningkatkan ketepatan prediksi hasil panen di sektor pertanian

3. HASIL DAN PEMBAHASAN

Bagian ini menyajikan hasil evaluasi dari implementasi algoritma *Random Forest Regressor* dan *Decision Tree Regressor* dalam sistem prediksi hasil panen. Hasil yang diperoleh dianalisis untuk menilai akurasi, stabilitas, dan efektivitas masing-masing model berdasarkan metrik evaluasi yang telah ditentukan. Pembahasan dilakukan untuk menginterpretasikan temuan dan menjawab rumusan masalah yang telah disusun sebelumnya.

3.1. Hasil Evaluasi Model dan Pemilihan Model

Evaluasi dilakukan untuk membandingkan performa algoritma *Random Forest Regressor* dan *Decision Tree Regressor* dalam sistem prediksi hasil panen. Penilaian dilakukan menggunakan metrik utama *Mean Absolute Error (MAE)*, yang mengukur rata-rata kesalahan absolut antara nilai prediksi dan nilai aktual.

Tabel 3.1 Hasil Evaluasi Model

Model	MAE
<i>Random Forest Regressor</i>	0.48
<i>Decision Tree Regressor</i>	0.43

Dari hasil evaluasi pada Tabel 3.1, terlihat bahwa *Decision Tree Regressor* memiliki nilai *MAE* yang lebih rendah, menunjukkan tingkat kesalahan prediksi yang lebih kecil dibandingkan *Random Forest Regressor*. Artinya, model ini lebih akurat dalam menghasilkan prediksi hasil panen berdasarkan data yang digunakan.

Selain itu, secara implementatif, *Decision Tree* memiliki keunggulan dalam hal kemudahan interpretasi karena struktur modelnya yang sederhana dan transparan. Berdasarkan evaluasi kuantitatif serta pertimbangan teknis dan fungsional, *Decision Tree Regressor* dipilih sebagai model terbaik dalam sistem prediksi hasil panen yang dikembangkan dalam penelitian ini.

3.2. Analisis Feature Importance

Untuk memahami kontribusi masing-masing fitur terhadap hasil prediksi, dilakukan analisis *feature importance* pada kedua model. Tabel 3.2 menyajikan nilai *feature importance* dari algoritma *Random Forest Regressor* dan *Decision Tree Regressor*

Tabel 3.2 Nilai Feature Importance pada Masing-Masing Model

No.	Feature	Random Forest	Decision Tree
1.	Rainfall mm	0.6374	0.6358
2.	Fertilizer Used	0.2013	0.2018
3.	Irrigation Used	0.1246	0.1262
4.	Crop	0.0142	0.0141
5.	Soil Type	0.014	0.014
6.	Weather Condition	0.0085	0.0082

Hasil pada Tabel 3.2 menunjukkan bahwa *Rainfall mm* merupakan fitur paling dominan pada kedua model, dengan nilai lebih dari 63%, menandakan pengaruh besar curah hujan terhadap hasil panen. Hal ini konsisten dengan pengetahuan agronomis bahwa curah hujan sangat menentukan pertumbuhan tanaman dan produktivitas lahan.

Fitur *Fertilizer Used* dan *Irrigation Used* memiliki kontribusi sedang, sekitar 12–20%, yang menunjukkan peran pentingnya dalam mendukung hasil panen, meskipun tidak sebesar curah hujan. Sementara itu, fitur lain seperti *Crop*, *Soil Type*, dan *Weather Condition* memberikan kontribusi yang relatif kecil (di bawah 2%), kemungkinan karena kurangnya variasi data atau korelasi yang rendah dengan target.

4. KESIMPULAN

Berdasarkan hasil penelitian mengenai perbandingan algoritma *Random Forest Regressor* dan *Decision Tree Regressor* untuk prediksi hasil panen, dapat disimpulkan bahwa *Decision Tree Regressor* merupakan model yang paling optimal untuk dataset yang digunakan. Hal ini dibuktikan dengan nilai *Mean Absolute Error (MAE)* yang lebih rendah, yaitu 0,43, dibandingkan 0,48 pada *Random Forest Regressor*, yang menunjukkan tingkat akurasi prediksi yang lebih tinggi. Selain itu, hasil analisis *feature*

importance dari kedua model menunjukkan bahwa curah hujan (*Rainfall_mm*) merupakan faktor paling berpengaruh terhadap hasil panen, diikuti oleh penggunaan pupuk dan sistem irigasi. Temuan ini memperkuat pentingnya ketiga faktor tersebut dalam mendukung keberhasilan produksi pertanian. Dengan demikian, Decision Tree Regressor tidak hanya unggul secara metrik evaluasi, tetapi juga memberikan interpretasi yang jelas terhadap faktor-faktor penentu hasil panen.

DAFTAR PUSTAKA

- [1] B. Pusat Statistik, *Analisis Produktivitas Jagung Dan Kedelai Di Indonesia The Analysis Of Maize And Soybean Yield In Indonesia (The Result Of Crop-Cutting Survey)*, vol. 4. Badan Pusat Statistik, 2024.
- [2] Badan Pusat Statistik, *potensi-pertanian-indonesia-peta-baru-pertanian-berkelanjutan*. Indonesia, 2023.
- [3] S. Si. Widya Khonik Zuraina, S. K. Eko Pudjianto, S. Kom. Asep Udin, and S. Si. M. Stat. Neny Kurniawati, *STATISTIK PERKEBUNAN UNGGULAN NASIONAL 2021-2023*. Jakarta: Sekretariat Direktorat Jenderal Perkebunan, 2023.
- [4] K. Pangan Berkelanjutan, F. Pertanian, U. Lampung Jl Sumantri Brojonegoro No, and B. Lampung, "Tantangan Sektor Pertanian dalam Memenuhi Challenges of The Agricultural Sector in Meeting the Needs for Sustainable Food Muhammad Ibnu," 2024. [Online]. Available: <http://>
- [5] D. Swain, S. Lakum, S. Patel, P. Patro, and Jatin, "An Efficient Crop Yield Prediction System Using Machine Learning," *EAI Endorsed Transactions on Internet of Things*, vol. 10, pp. 1–5, Dec. 2024, doi: 10.4108/eetiot.5333.
- [6] N. Nur, F. Wajidi, S. Sulhayanti, and W. Wildayani, "Implementasi Algoritma Random Forest Regression untuk Memprediksi Hasil Panen Padi di Desa Minanga," *Jurnal Komputer Terapan*, vol. 9, no. 1, pp. 58–64, Jun. 2023, doi: 10.35143/jkt.v9i1.5917.
- [7] D. Muallifah, W. Fadila, and R. Firdaus, "Teknik SMOTE untuk Mengatasi Imbalance Data pada Deteksi Penyakit Stroke Menggunakan Algoritma Random Forest," *Jurnal CoSciTech (Computer Science and Information Technology)*, vol. 3, no. 2, pp. 107–113, Aug. 2022, doi: 10.37859/coscitech.v3i2.3912.
- [8] Fitri Handayani and Reny Medikawati Taufiq, "Komparasi Algoritma Menggunakan Teknik Smote Dalam Melakukan Klasifikasi Penyakit Stroke Otak," *Jurnal CoSciTech (Computer Science and Information Technology)*, vol. 5, no. 2, pp. 367–372, Aug. 2024, doi: 10.37859/coscitech.v5i2.7439.
- [9] F. Indra Kurniadi, D. Satyananda, E. Santika, and P. Dwi Larasati, "Multi-Output Regression untuk Melakukan Prediksi Luas Wilayah, Kualitas Padi dan Produksi Padi pada Pulau Jawa," *Jurnal Sistem Komputer & Kecerdasan Buatan(SISKOM-KB)*, vol. 5, pp. 1–6, 2022, doi: <https://doi.org/10.47970/siskom-kb.v5i2.269>.
- [10] A. I. Ramadhan, N. Almajid, and D. Y. Ginting, "Penerapan Decision Tree Regression dalam Memprediksi Harga Rumah di Provinsi Jawa Barat," *Jejaring Penelitian dan Pengabdian Masyarakat (JPPM)*, vol. 1, no. 3, pp. 111–115, 2024.
- [11] X. Huang, "Predictive Models: Regression, Decision Trees, and Clustering," *Applied and Computational Engineering*, vol. 79, no. 1, pp. 124–133, Jul. 2024, doi: 10.54254/2755-2721/79/20241551.
- [12] P. Dumre, S. Bhattarai, and H. K. Shashikala, "Optimizing Linear Regression Models: A Comparative Study of Error Metrics," in *Proceedings - 4th International Conference on Technological Advancements in Computational Sciences, ICTACS 2024*, Institute of Electrical and Electronics Engineers Inc., 2024, pp. 1856–1861. doi: 10.1109/ICTACS62700.2024.10840719.