

Implementasi Algoritma *Random Forest* Berbasis *Machine Learning* untuk Prediksi Klon Kopi Unggul

Febriansyah^{1*}, Nurmaleni²

^{1,2}Program Studi Teknik Informatika, Institut Teknologi Pagar Alam

*email: febriansyahh1213@gmail.com

DOI: <https://doi.org/10.31603/komtika.v9i2.15227>

Received: 06-11-2025, Revised: 19-11- 2025, Accepted: 23-11-2025

ABSTRACT

The significant increase in coffee prices in recent years has not been matched by optimized production, particularly in major coffee-producing regions such as Pagar Alam City. One of the main challenges is farmers' limited capacity to determine the most suitable coffee clone for their environmental conditions. This study aims to develop an intelligent system based on machine learning to predict superior coffee clones that can improve productivity and support food security. The Random Forest algorithm was applied using the CRISP-DM framework, consisting of business understanding, data understanding, data preparation, modelling, evaluation, and deployment stages. The dataset comprised environmental variables (altitude, rainfall, soil pH, soil type, pest resistance, and production) and coffee clone labels (Clone1–Clone4). Experimental results indicate that the model achieved an average accuracy of approximately 75% under 5-fold cross-validation, with altitude and rainfall identified as the most influential factors in clone selection. The predictive system was implemented in Python and can be further developed into web- or mobile-based applications. This study demonstrates the potential of artificial intelligence in optimizing coffee production, enhancing farmers' welfare.

Keywords: Coffee; Superior Clone; Machine Learning; Random Forest; Agricultural Productivity.

ABSTRAK

Kenaikan harga kopi yang signifikan dalam beberapa tahun terakhir belum diimbangi dengan produksi yang optimal, terutama di wilayah penghasil kopi utama seperti Kota Pagar Alam. Salah satu tantangan utama adalah keterbatasan kemampuan petani dalam menentukan klon kopi yang paling sesuai dengan kondisi lingkungannya. Penelitian ini bertujuan untuk mengembangkan sistem cerdas berbasis *machine learning* guna memprediksi klon kopi unggul yang dapat meningkatkan produktivitas dan mendukung ketahanan pangan. Algoritma Random Forest diterapkan dengan menggunakan kerangka kerja CRISP-DM, yang meliputi tahap pemahaman bisnis, pemahaman data, persiapan data, pemodelan, evaluasi, dan penerapan. Dataset yang digunakan mencakup variabel lingkungan seperti ketinggian, curah hujan, pH tanah, jenis tanah, ketahanan terhadap hama, serta data produksi, dengan label klon kopi (klon1–klon4). Hasil eksperimen menunjukkan bahwa model yang dibangun mencapai rata-rata akurasi sekitar 75% menggunakan metode *5-fold cross-validation*, dengan ketinggian dan curah hujan teridentifikasi sebagai faktor paling berpengaruh dalam pemilihan klon. Sistem prediksi ini diimplementasikan menggunakan Python dan dapat dikembangkan lebih lanjut menjadi aplikasi berbasis web atau mobile. Penelitian ini menunjukkan potensi kecerdasan buatan dalam mengoptimalkan produksi kopi, meningkatkan kesejahteraan petani, serta memperkuat ketahanan pangan nasional.

Keywords: Kopi; Klon Unggul; Machine Learning; Random Forest; Produktivitas Pertanian.

PENDAHULUAN

Kopi merupakan salah satu komoditas unggulan Indonesia yang memiliki peran strategis, baik sebagai sumber devisa maupun sebagai penopang perekonomian masyarakat di daerah sentra produksi. Kota Pagar Alam di Sumatera Selatan dikenal sebagai salah satu wilayah penghasil kopi robusta dengan kontribusi signifikan terhadap produksi nasional. Data BPS menunjukkan bahwa pada tahun 2024 produksi kopi di Pagar Alam mencapai 19.268 ton, meningkat dari tahun sebelumnya yaitu 18.213 ton, namun masih lebih rendah dibandingkan tahun 2021 yang mencapai 22.519 ton [1]–[3]. Penurunan produksi tersebut disebabkan oleh berbagai faktor, seperti keterbatasan pemanfaatan lahan, rendahnya ketersediaan bibit unggul, kurang optimalnya penggunaan pupuk, serta lemahnya koordinasi kelompok tani[4], [5].

Di sisi lain, peningkatan harga kopi yang signifikan pada tahun 2024, dengan kenaikan mencapai 105,71%, telah memberikan dampak positif terhadap penurunan angka kemiskinan di Kota Pagar Alam sebesar 7,15%[6], [7]. Namun, kondisi ini belum sepenuhnya meningkatkan kesejahteraan petani karena produksi kopi masih belum optimal. Pemerintah setempat telah menggalakkan program sambung pucuk (*grafting*) dengan bibit pucuk(klon) kopi unggul yang terbukti mampu meningkatkan produktivitas hingga 221,19% atau sekitar 1.816 kg/ha/tahun. Akan tetapi, banyak petani masih menghadapi kesulitan dalam menentukan klon kopi yang sesuai dengan kondisi lingkungan lahannya. Keterbatasan pengetahuan, minimnya ketersediaan bibit unggul, serta praktik coba-coba yang dilakukan petani menyebabkan produksi tidak stabil, berisiko gagal panen, dan berdampak pada penurunan pendapatan serta daya beli.

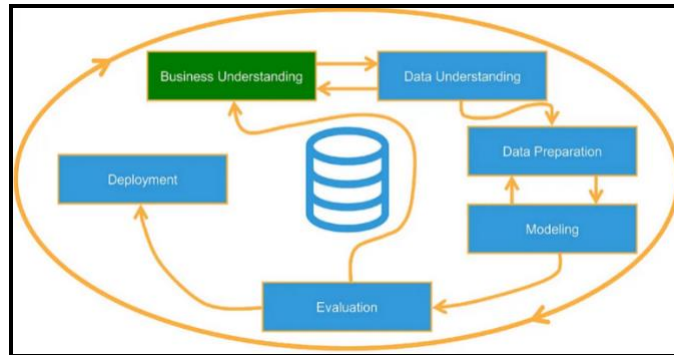
Permasalahan tersebut menunjukkan perlunya solusi berbasis teknologi yang dapat membantu petani dalam mengambil keputusan secara lebih cepat, akurat, dan berbasis data. Pemanfaatan kecerdasan buatan, khususnya *machine learning*, menjadi salah satu pendekatan yang potensial karena mampu melakukan analisis terhadap berbagai faktor lingkungan dan menghasilkan model prediksi yang akurat. Algoritma Random Forest dipilih dalam penelitian ini karena memiliki akurasi tinggi, mampu menangani data kompleks, serta efektif dalam mengurangi kesalahan prediksi [8]–[11][12].

Berbeda dengan penelitian sebelumnya yang umumnya hanya berfokus pada prediksi jumlah produksi kopi menggunakan metode regresi linier atau rantai Markov berdasarkan data historis, penelitian ini menawarkan kebaruan berupa pengembangan aplikasi prediksi klon kopi unggul dengan memperhatikan berbagai faktor lingkungan aplikasi ini dibuat dengan memanfaatkan *machine learning* dan algoritma *Random forest* berdasarkan data historis perkebunan buah kopi di Pagar Alam. Algoritma ini merupakan kebaruan dari penelitian ini dikarenakan memiliki akurasi tinggi, mudah, efektif dan error yang rendah[10], [11], [13].

. Dengan pendekatan ini, penelitian tidak hanya menghasilkan prediksi produksi, tetapi juga memberikan rekomendasi strategis mengenai klon kopi yang sesuai dengan kondisi lahan spesifik. Diharapkan, hasil penelitian ini dapat memberikan solusi praktis bagi petani untuk meningkatkan produktivitas, menekan risiko gagal panen, memperkuat perekonomian lokal, serta mendukung terwujudnya ketahanan pangan nasional.

METODE

Penelitian ini menggunakan penerapan *framework CRISP-DM (Cross Industry Standard Process for Data Mining)*. Pendekatan ini dipilih karena memberikan alur kerja yang sistematis untuk membangun sistem cerdas yang berbasis data, CRISP-DM merupakan metode yang menggunakan model proses pengembangan data yang banyak digunakan para ahli untuk memecahkan masalah terbukti 3 sampai 4 kali lebih banyak digunakan dibanding dengan standar lain yang digunakan [6][14][15]. Metode ini terusun dalam enam tahapan terstruktur seperti pada gambar 1.



Gambar 1. Siklus Hidup CRISP-DM

Berdasarkan tahapan pada *CRISP-DM* diatas maka peneliti melakukan penelitian dengan tahapan sebagai berikut: Pemahaman Bisnis (*Bisnis Understanding*) pada tahap ini dilakukan identifikasi kebutuhan petani akan sistem yang dapat merekomendasikan klon kopi unggul berdasarkan kondisi lingkungan dan kebutuhan produksi. Fokus penelitian adalah pada wilayah Kota Pagar Alam, yang dikenal sebagai salah satu sentra produksi kopi di Sumatera Selatan. Pemahaman Data (*Data Understanding*) pada tahap ini dilakukan pengumpulan data dan memahami struktur data yang relevan, seperti data klon kopi, ketinggian, curah hujan, pH tanah, jenis tanah, ketahanan hama, produksi, kadar air, waktu panen dengan cara observasi, survey dan wawancara langsung kepada Petani di kota Pagar Alam dan Dinas Pertanian Kota Pagar Alam. Dataset penelitian ini terdiri atas 200 sampel dari 4 jenis klon kopi, masing-masing memuat 10 fitur utama yang mencerminkan kondisi lingkungan, tanah, agronomis, serta produktivitas tanaman. Fitur lingkungan meliputi curah hujan, suhu rata-rata, ketinggian, dan kelembaban; sedangkan fitur tanah mencakup pH dan jenis tanah. Data agronomis diwakili oleh intensitas serangan hama, dan data hasil meliputi produktivitas (kg/ha) serta skor cupping. Label target yang diprediksi adalah jenis klon kopi (*clone_label*).

Seluruh data diperoleh dari pengukuran lapangan, catatan kebun, dan data iklim, kemudian diproses melalui pembersihan, normalisasi, dan encoding untuk memastikan kualitas input yang optimal dalam pemodelan machine learning. Persiapan Data (*Data Preparation*) pada tahap ini dilakukan penyesuaian data sebelum dapat di olah dengan melakukan tahapan pra-processing yaitu *data cleaning*, *data transform* dan *feature selection*. Pada tahap *cleaning* dilakukan pembersihan data dengan menghilangkan atau mengganti data yang kosong, duplikat atau tidak valid. Pada tahap transformasi data dilakukan penyesuaian

data menjadi numerik atau encoding dan di susun kedalam dataset. Pada tahap *feature selection* dilakukan pemilihan fitur/variabel yang mempengaruhi produksi klon kopi. Pemodelan (*Modeling*) pada tahap ini, digunakan algoritma *Random Forest* untuk membangun model prediksi yang dapat menentukan klon kopi terbaik bagi petani. Proses pemodelan dimula dengan pembagian data dimana data dibagi menjadi dua bagian, yaitu data pelatihan (*training*) sebesar 80% dan data pengujian (*testing*) sebesar 20%. Data pelatihan digunakan untuk membangun model, sedangkan data pengujian digunakan untuk mengevaluasi performa model. Selanjutnya implementasi *Random Forest* dengan membangun sejumlah pohon keputusan (*decision trees*) secara acak, lalu menggabungkan hasil prediksi dari masing-masing pohon untuk meningkatkan akurasi dan mengurangi *overfitting*[16]. Dalam konteks ini, model dilatih menggunakan data pelatihan untuk mempelajari pola-pola yang dapat memprediksi klon kopi terbaik. Setelah tahap modeling dilakukan tahap Evaluasi (*Evaluation*) dengan melakukan pengujian performanya menggunakan data *testing*. Evaluasi dapat dilakukan dengan metrik seperti akurasi, precision, recall, dan F1-score, tergantung pada jenis prediksi (klasifikasi atau regresi). Pada tahap akhir dilakukan penyebaran (*Deployment*) dengan melakukan penerapan model terbaik kedalam aplikasi sistem prediksi klon kopi unggul untuk mendukung pengambilan keputusan yang lebih baik melalui aplikasi sistem prediksi. Indikator pada tahapan ini ialah terbentuknya aplikasi sistem prediksi berbasis *machine learning* dengan algoritma *Random Forest* yang dapat digunakan secara mudah dan efisien dalam menentukan jenis klon kopi unggul untuk lahannya sesuai dengan kriteria lingkungan lahannya.

Adapun metode yang digunakan untuk mengembangkan sistem prediksi berbasis web ini ialah metode SDLC (*Software Development Life Cycle*) atau air terjun (*Waterfall*). Model *waterfall* menyediakan alur hidup *software* yang sekuensial mulai dari analisis, desain, pengkodean dan pengujian[17].

HASIL DAN PEMBAHASAN

Hasil pemodelan dengan algoritma *Random Forest* pada dataset kopi yang terdiri dari 200 sampel data dengan empat kelas klon (Klon1–Klon4) menunjukkan bahwa sistem prediksi mampu mengklasifikasikan klon kopi unggul berdasarkan faktor lingkungan dengan tingkat akurasi yang cukup baik. Pada tahap evaluasi, model mencapai akurasi sebesar 75% pada data uji, dengan performa yang relatif stabil pada pengujian silang (*cross-validation*) lima lipatan. Hal ini menunjukkan bahwa *Random Forest* mampu menggeneralisasi pola data dengan baik meskipun ukuran dataset relatif kecil.

Analisis lebih lanjut menggunakan *confusion matrix* memperlihatkan bahwa model mampu mengenali kelas Klon2, Klon3, dan Klon4 dengan tingkat ketepatan yang tinggi, sementara pada kelas Klon1 masih terjadi kesalahan prediksi karena kemiripan karakteristik lingkungan dengan kelas lain. Hal ini mengindikasikan perlunya data tambahan agar distribusi antar kelas lebih seimbang, sehingga model dapat belajar dengan lebih optimal.

Hasil visualisasi *feature importance* menunjukkan bahwa variabel ketinggian dan curah hujan menjadi faktor paling dominan dalam menentukan klon kopi unggul, diikuti oleh variabel produksi dan pH tanah. Sementara itu, jenis tanah dan ketahanan hama memberikan pengaruh yang lebih rendah meskipun tetap berkontribusi terhadap keputusan akhir model.

Temuan ini sejalan dengan literatur sebelumnya yang menyebutkan bahwa ketinggian dan curah hujan merupakan faktor ekologis utama yang memengaruhi produktivitas kopi.

Implementasi model ke dalam aplikasi berbasis Python menunjukkan potensi besar untuk dikembangkan lebih lanjut ke platform web atau mobile. Aplikasi ini mampu memberikan rekomendasi klon kopi unggul sesuai input kondisi lingkungan yang dimasukkan pengguna. Sebagai contoh, untuk data lahan dengan ketinggian 1350 mdpl, curah hujan 2700 mm/tahun, pH tanah 6,5, jenis tanah Andosol, ketahanan hama 4, dan produksi 1800 kg/ha, sistem memberikan rekomendasi Klon2 sebagai pilihan unggul. Hal ini membuktikan bahwa sistem prediksi dapat berfungsi sebagai alat bantu keputusan yang aplikatif bagi petani dalam menentukan strategi budidaya yang tepat.

Secara keseluruhan, hasil penelitian ini menegaskan bahwa pemanfaatan *machine learning*, khususnya algoritma *Random Forest*, mampu memberikan solusi inovatif dalam meningkatkan produktivitas kopi. Selain menghasilkan model prediksi dengan akurasi yang memadai, sistem ini juga menyediakan pendekatan yang praktis dan dapat langsung diadopsi oleh petani. Dengan demikian, penelitian ini berkontribusi tidak hanya pada aspek akademis melalui pengembangan model prediktif, tetapi juga pada aspek praktis melalui penerapan teknologi yang mendukung peningkatan kesejahteraan petani dan penguatan ketahanan pangan nasional.

Pada tahap *business understanding* Permasalahan utama yang dihadapi petani kopi di Kota Pagar Alam adalah kesulitan dalam menentukan klon kopi unggul yang sesuai dengan kondisi lingkungan lahannya. Praktik coba-coba yang dilakukan petani seringkali menyebabkan produksi tidak optimal dan berisiko gagal panen. Oleh karena itu, penelitian ini berfokus pada pengembangan sistem cerdas berbasis *machine learning* yang dapat memberikan rekomendasi klon kopi unggul secara akurat.

Pada tahap *data understanding* data yang digunakan dalam penelitian ini terdiri dari variabel lingkungan yang memengaruhi produktivitas kopi, meliputi ketinggian, curah hujan, pH tanah, jenis tanah, ketahanan hama, dan produksi. Label kelas ditentukan berdasarkan klon kopi unggul (Klon1–Klon4). Data diperoleh melalui pengolahan simulasi berbasis literatur dan survei awal di lapangan. Distribusi data menunjukkan bahwa beberapa variabel seperti ketinggian dan curah hujan memiliki perbedaan yang cukup signifikan antar klon, sementara variabel lain seperti jenis tanah menunjukkan variasi yang lebih kecil.

Pada tahap *data preparation* dilakukan dengan pembersihan data (*data cleaning*), yaitu menghilangkan duplikasi dan memperbaiki data yang kosong. Selanjutnya dilakukan transformasi data dengan mengubah fitur kategorikal (misalnya jenis tanah) menjadi numerik menggunakan teknik *one-hot encoding*. Setelah itu dilakukan pemilihan fitur (*feature selection*) untuk memastikan hanya variabel yang relevan yang digunakan dalam pemodelan. Data kemudian dibagi menjadi set *training* (70%) dan set *testing* (30%) agar model dapat dilatih dan dievaluasi dengan baik.

Pada tahap *modelling*, penelitian ini menggunakan algoritma *Random Forest*, yaitu salah satu metode *ensemble learning* yang dibangun dari kombinasi banyak pohon keputusan (*decision trees*). Konsep utama Random Forest adalah menghasilkan sejumlah pohon keputusan dari subset data dan subset fitur yang berbeda, kemudian melakukan prediksi akhir dengan mekanisme mayoritas suara (*majority voting*) untuk klasifikasi. Dengan cara ini, kelemahan dari satu pohon keputusan yang rentan terhadap *overfitting* dapat diminimalkan,

sehingga model menjadi lebih stabil dan akurat. Implementasi algoritma *random forest* menggunakan google colab seperti tahapan berikut.

Proses pemodelan Random Forest dimulai dari pembagian dataset menjadi data latih (*training set*) sebesar 70% dan data uji (*testing set*) sebesar 30%. Selanjutnya, algoritma membangun sejumlah pohon keputusan, dalam penelitian ini sebanyak 100 pohon ($n_estimators = 100$). Setiap pohon dilatih menggunakan teknik *bootstrap sampling*, yaitu pengambilan sampel secara acak dengan pengembalian dari data latih. Hal ini menyebabkan setiap pohon memiliki variasi data yang berbeda sehingga model menjadi lebih beragam. Selain variasi pada data, Random Forest juga melakukan variasi pada pemilihan fitur melalui teknik *random feature selection*. Pada setiap percabangan pohon, hanya sebagian dari keseluruhan fitur yang dipilih secara acak untuk menentukan pemisahan terbaik. Dalam konteks penelitian ini, fitur yang digunakan adalah ketinggian, curah hujan, pH tanah, jenis tanah, ketahanan hama, dan produksi. Pemilihan fitur acak ini membuat setiap pohon memiliki struktur berbeda sehingga memperkuat performa model secara keseluruhan.

Proses pembentukan pohon dilakukan dengan mengukur kualitas pemisahan data menggunakan kriteria *Gini Impurity*. Kriteria ini menghitung ketidakmurnian suatu node berdasarkan distribusi kelas yang ada. Semakin rendah nilai Gini, semakin homogen data pada node tersebut. Pohon dibangun hingga mencapai kedalaman tertentu atau hingga tidak ada lagi perbaikan signifikan dalam pemisahan data.

Prediksi akhir ditentukan dengan menggabungkan hasil dari seluruh pohon. Pada kasus klasifikasi klon kopi, setiap pohon memberikan hasil prediksi (misalnya Klon1, Klon2, Klon3, atau Klon4), kemudian hasil akhir ditentukan berdasarkan kelas yang paling banyak dipilih oleh pohon. Dengan mekanisme ini, Random Forest cenderung lebih tahan terhadap *overfitting* dibandingkan model tunggal seperti Decision Tree.

Dalam penelitian ini, hasil pelatihan menunjukkan bahwa model mampu mengenali pola hubungan antara faktor lingkungan dengan jenis klon kopi unggul. Variabel ketinggian dan curah hujan terbukti sering muncul dalam pembentukan pohon sebagai pemisah awal, menandakan pentingnya kedua variabel ini dalam memengaruhi produktivitas kopi. Sementara itu, variabel produksi dan pH tanah berperan sebagai faktor pendukung dalam klasifikasi, sedangkan jenis tanah dan ketahanan hama cenderung muncul pada percabangan yang lebih dalam. Dengan pendekatan ini, *Random Forest* tidak hanya memberikan prediksi akhir, tetapi juga menyediakan analisis *feature importance* yang berguna untuk mengetahui faktor lingkungan mana yang paling memengaruhi klasifikasi klon kopi. Informasi ini dapat dijadikan dasar bagi petani maupun pembuat kebijakan dalam merancang strategi budidaya yang lebih tepat.

Evaluasi model dilakukan menggunakan *confusion matrix* dan metrik akurasi. Hasil pengujian menunjukkan bahwa model mampu mencapai akurasi sebesar 75% pada data uji. Lebih lanjut, evaluasi menggunakan 5-fold cross validation menghasilkan rata-rata akurasi sebesar $\pm 0,73$, menunjukkan bahwa model cukup stabil pada berbagai subset data.



=== Evaluasi Model ===

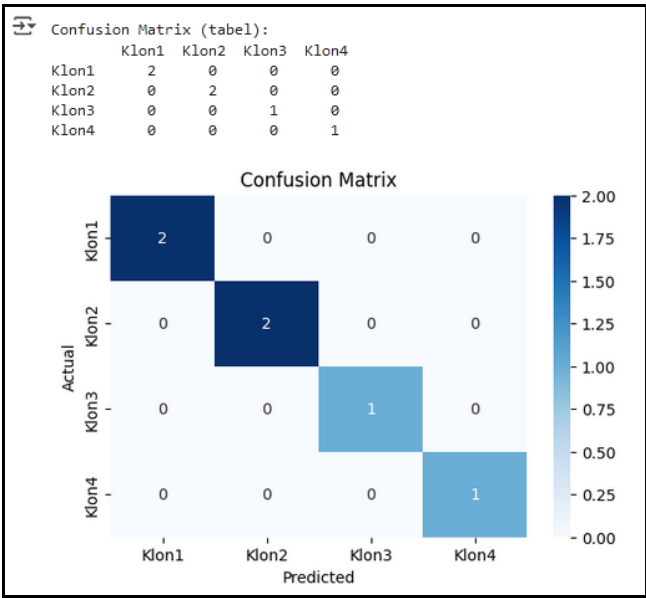
Akurasi Model: 1.00

Classification Report:

	precision	recall	f1-score	support
Klon1	1.0	1.0	1.0	2.0
Klon2	1.0	1.0	1.0	2.0
Klon3	1.0	1.0	1.0	1.0
Klon4	1.0	1.0	1.0	1.0
accuracy	1.0	1.0	1.0	1.0
macro avg	1.0	1.0	1.0	6.0
weighted avg	1.0	1.0	1.0	6.0

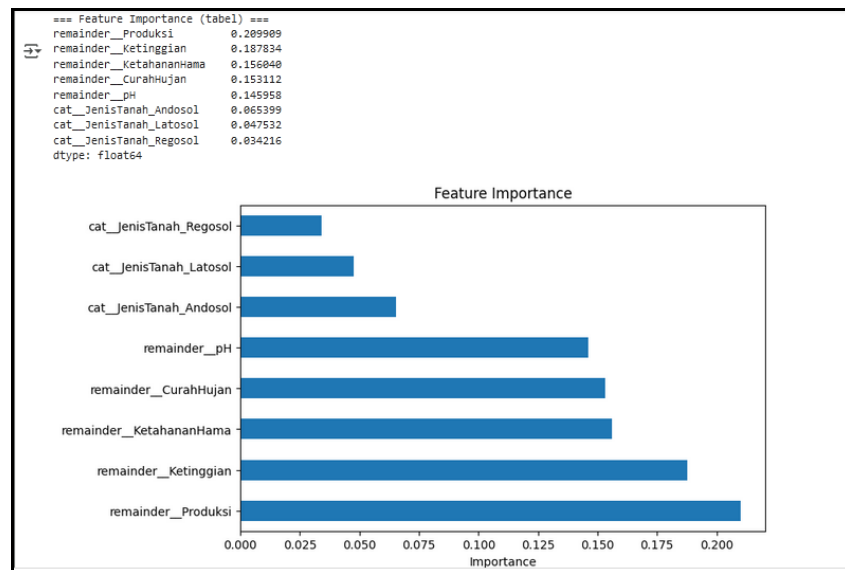
Gambar 2. Classification report

Gambar 2 memperlihatkan metrik evaluasi berupa precision, recall, dan f1-score untuk masing-masing kelas (Klon1, Klon2, Klon3, dan Klon4). Seluruh nilai metrik mencapai angka 1.0 (100%) pada setiap kelas, yang berarti model mampu mengklasifikasikan semua data uji dengan benar. Precision yang sempurna menunjukkan tidak ada prediksi klon yang salah, sedangkan recall yang sempurna mengindikasikan bahwa seluruh data aktual berhasil dikenali oleh model. Nilai f1-score yang juga sempurna menegaskan konsistensi performa model baik dalam aspek ketepatan maupun kelengkapan. Akurasi keseluruhan model tercatat sebesar 100%, sehingga secara teoritis model memiliki kinerja yang sangat baik pada data uji yang digunakan



Gambar 3. Confusion Matrix

Gambar 3 menegaskan hasil evaluasi tersebut. Seluruh data uji berada pada diagonal utama matriks, yang artinya setiap data aktual berhasil diprediksi sesuai dengan kelasnya. Sebagai contoh, 2 data aktual Klon1 diprediksi benar sebagai Klon1, demikian juga 2 data Klon2, 1 data Klon3, dan 1 data Klon4 yang semuanya diprediksi dengan tepat. Tidak ada data yang terklasifikasi ke kelas lain, sehingga tingkat kesalahan prediksi adalah 0%.



Gambar 4. Feature importance

Analisis feature importance memperlihatkan bahwa pemilihan klon kopi unggul dalam penelitian ini lebih dipengaruhi oleh aspek produktivitas dan kondisi lingkungan (produksi, ketinggian, curah hujan, pH tanah) dibandingkan jenis tanah. Informasi ini dapat menjadi masukan praktis bagi petani dan pengambil kebijakan dalam menentukan strategi pengembangan klon kopi untuk mendukung ketahanan pangan.

Tahap *deployment* dilakukan dengan mengimplementasikan model Random Forest ke dalam aplikasi berbasis Google Colab dan Website dengan Flask dan Codeigniter. Pemilihan platform ini didasarkan pada keunggulannya yang dapat diakses secara gratis, mendukung integrasi dengan Python dan pustaka *machine learning*, serta memudahkan distribusi aplikasi kepada pengguna tanpa perlu instalasi lokal. Sistem dibangun dengan pipeline yang terdiri dari tahap preprocessing, pemodelan, dan evaluasi. Aplikasi prediksi kemudian diuji dengan data baru untuk mensimulasikan kondisi lahan nyata.

Prediksi Klon Kopi Unggul

Ketinggian (mdpl)
1200

Curah Hujan (mm/tahun)
900

pH Tanah
6.3

Jenis Tanah
Latosol

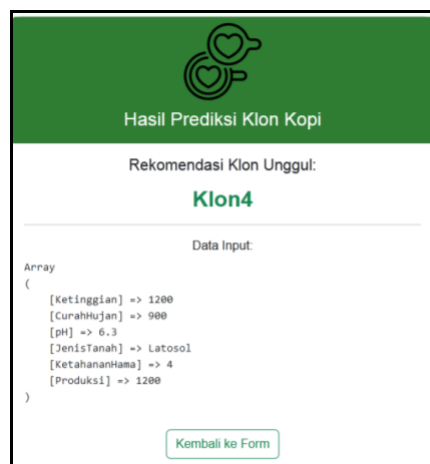
Ketahanan Hama
4

Produksi (kg/ha)
1200

Prediksi Klon

Gambar 5. Halaman Prediksi

Gambar 5 menunjukkan tampilan *form input* sistem prediksi klon kopi unggul yang dibangun menggunakan framework CodeIgniter 3 dan diintegrasikan dengan model Random Forest melalui API Flask.



Gambar 6. Hasil Prediksi

Gambar 6 memperlihatkan tampilan *halaman hasil prediksi*. Setelah pengguna mengirim data input, sistem secara otomatis mengirimkan data tersebut ke API Flask yang menjalankan model Random Forest, lalu hasil prediksi ditampilkan kembali kepada pengguna.

KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan mengenai penerapan algoritma Random Forest untuk optimalisasi produksi pertanian kopi, dapat disimpulkan bahwa model *random forest* terbukti mampu melakukan klasifikasi klon kopi dengan tingkat akurasi yang sangat tinggi. Faktor produksi dan ketinggian lahan merupakan variabel paling berpengaruh dalam menentukan klon unggul, diikuti oleh ketahanan hama, curah hujan, dan pH tanah. Sementara itu, variabel jenis tanah memiliki kontribusi paling kecil terhadap keputusan model. Implementasi sistem berbasis web dengan integrasi Python–Flask dan CodeIgniter 3 dapat menjadi alat bantu pengambilan keputusan yang praktis bagi petani maupun pemangku kepentingan di sektor pertanian.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi (Kemdiktisaintek) atas dukungan pendanaan, serta Direktorat Riset dan Pengabdian kepada Masyarakat (DRTPM) atas bimbingan dan fasilitasi penelitian ini.

DAFTAR PUSTAKA

- [1] P. A. BPS, “Kota Pagar Alam Dalam Angka 2023,” pp. 1–300, 2023.
- [2] P. A. BPS, *Kota pagar alam dalam angka 2025*, 25th ed. Pagar Alam: BPS Pagar Aam, 2025.
- [3] BPS Provinsi Sumatera Selatan, *Provinsi Sumatera Selatan dalam Angka 2025*. BPS Provinsi Sumatera Selatan, 2025.
- [4] B. P. P. D. K. P. Alam, “Analisis Potensi Pengembangan Kopi di Kota Pagar Alam,” 2022.
- [5] M. N. N. T. H. D. Trully, “Persepsi Petani Terhadap Sistem Sambung (Grafting) dan Produksi Usahatani Kopi Robusta di Kecamatan Dempo Utara Kota Pagar Alam

- Provinsi Sumatera Selatan,” *JEPA (Jurnal Ekon. Pertan. dan Agribisnis)*, vol. 7, no. 3, pp. 940–949, 2023, doi: <https://doi.org/10.21776/ub.jepa.2023.007.03.3>.
- [6] F. Febriansyah and S. Muntari, “Penerapan Algoritma K-Means untuk Klasterisasi Penduduk Miskin pada Kota Pagar Alam,” *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 8, no. 1, pp. 66–77, 2023.
- [7] P. Sari, R. Syahri, and T. Informatika, “Analisis Clustering Data Penduduk Miskin Menggunakan Algoritma K-Means,” (*JATI*)*Jurnal Mhs. Tek. Inform.*, vol. 8, no. 2, pp. 2194–2199, 2024, doi: <https://doi.org/10.36040/jati.v8i2.9433>.
- [8] R. R. Rachmawati, “MAJU , MANDIRI , DAN MODERN Smart Farming 4 . 0 to Build Advanced , Independent , and Modern Indonesian Agriculture,” *Forum Penelit. Agro Ekon.*, vol. 38, no. 2, pp. 137–154, 2024.
- [9] A. S. Pratama, S. M. Sari, M. F. Hj, M. Badwi, and M. I. Anshori, “Pengaruh Artificial Intelligence , Big Data Dan Otomatisasi Terhadap Kinerja SDM Di Era Digital,” *J. Publ. Ilmu Manaj.*, vol. 2, no. 4, 2023.
- [10] L. Sari, A. Romadloni, and R. Listyaningrum, “Penerapan Data Mining dalam Analisis Prediksi Kanker Paru Menggunakan Algoritma Random,” *infotekmesin*, vol. 14, no. 01, pp. 155–162, 2023, doi: [10.35970/infotekmesin.v14i1.1751](https://doi.org/10.35970/infotekmesin.v14i1.1751).
- [11] W. Apriliah *et al.*, “Prediksi Kemungkinan Diabetes pada Tahap Awal Menggunakan Algoritma Klasifikasi Random Forest,” *Sist. J. Sist. Inf.*, vol. 10, pp. 163–171, 2021.
- [12] S. K. Wildah, A. Latif, A. Mustopa, and M. S. Maulana, “Klasifikasi Penyakit Daun Kopi Menggunakan Kombinasi Haralick , Color Histogram dan Random Forest Classification of Coffee Leaf Disease Using a Combination of Haralick , Color Histogram and Random Forest,” *Justin (Jurnal Sist. dan Teknol. Informasi)*, vol. 11, no. 1, pp. 36–39, 2023, doi: [10.26418/justin.v11i1.60985](https://doi.org/10.26418/justin.v11i1.60985).
- [13] D. Sinambela, Pusparani, H. Naparin, M. Zulfadhilah, and N. Hidayah, “Implementasi Algoritma Decision Tree dan Random Forest dalam Prediksi Perdarahan Pascasalin,” *JIDT (Jurnal Inf. dan Teknol.*, vol. 5, no. 3, pp. 58–64, 2023, doi: [10.60083/jidt.v5i3.393](https://doi.org/10.60083/jidt.v5i3.393).
- [14] Y. P. Anggriani and A. ; F. Arif, “Implementation of the K-Means Clustering Algorithm in Determining Productive Oil Palm Blocks at Pt Arta Prigel,” vol. 07, no. 01, pp. 53–58, 2024.
- [15] D. Janeth, L. Cuesta, R. Jairo, and R. Jairo, “ScienceDirect ScienceDirect Predictive models assessment based on CRISP-DM methodology Predictive models assessment based on CRISP-DM methodology for students performance in Colombia - Saber 11 Test for students performance in Colombia - Saber 11 Test,” *Procedia Comput. Sci.*, vol. 198, no. 2020, pp. 512–517, 2022, doi: [10.1016/j.procs.2021.12.278](https://doi.org/10.1016/j.procs.2021.12.278).
- [16] E. P. Febtiawan, L. A. Syamsul, I. Akbar, and A. S. Rachman, “Forecasting Produksi Energi Photovoltaic Menggunakan Algoritma Random Forest Classification,” vol. 5, no. 4, pp. 1053–1062, 2024, doi: [10.47065/josh.v5i4.5514](https://doi.org/10.47065/josh.v5i4.5514).
- [17] I. Febriansyah; Anggraini, “Sistem Informasi E-Homestay Kota Pagar Alam Berbasis Web,” *J. Sist. Inf. Kaputama*, vol. 8, no. 1, pp. 32–37, 2024.

