

**KLASIFIKASI MUSIM TANAM TANAMAN HORTIKULTURA
BERDASARKAN PARAMETER CUACA DAN TANAH
MENGUNAKAN ENSEMBLE LEARNING**

SKRIPSI

DISUSUN OLEH

DWI NISYATUL WARDAH

NPM. 2209020175



**PROGRAM STUDI TEKNOLOGI INFORMASI
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA**

MEDAN

2026

**KLASIFIKASI MUSIM TANAM TANAMAN HORTIKULTURA
BERDASARKAN PARAMETER CUACA DAN TANAH
MENGUNAKAN ENSEMBLE LEARNING**

SKRIPSI

**Diajukan sebagai salah satu syarat untuk memperoleh gelar Sarjana
Komputer (S.Kom) dalam Program Studi Teknologi Informasi pada
Fakultas Ilmu Komputer dan Teknologi Informasi, Universitas
Muhammadiyah Sumatera Utara**

Dwi Nisyatul Wardah

NPM. 2209020175

**PROGRAM STUDI TEKNOLOGI INFORMASI
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA**

MEDAN

2026

LEMBAR PENGESAHAN

Judul Skripsi : KLASIFIKASI MUSIM TANAM TANAMAN
HORTIKULTURA BERDASARKAN PARAMETER
CUACA DAN TANAH MENGGUNAKAN ENSEMBLE
LEARNING
Nama Mahasiswa : DWI NISYATUL WARDAH
NPM : 2209020175
Program Studi : TEKNOLOGI INFORMASI

Menyetujui
Komisi Pembimbing



(Dr. Firaumi Rizky, S.Kom., M.Kom)
NIDN. 0116079201

**Ketua Program Studi
Teknologi Informasi**



(Fatma Sari Hutagalung, S.Kom, M.Kom)
NIDN. 0117019301

Dekan



(Dr. Al-Khowarizmi, S.Kom., M.Kom.)
NIDN. 0127099201

PERNYATAAN ORISINALITAS

KLASIFIKASI MUSIM TANAM TANAMAN HORTIKULTURA BERDASARKAN PARAMETER CUACA DAN TANAH MENGUNAKAN ENSEMBLE LEARNING

SKRIPSI

Saya menyatakan bahwa karya tulis ini adalah hasil karya sendiri, kecuali beberapa kutipan dan ringkasan yang masing-masing disebutkan sumbernya.

Medan, April 2026

Yang membuat pernyataan



Dwi Nisyatul Wardah

NPM. 2209020175

**PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN
AKADEMIS**

Sebagai sivitas akademika Universitas Muhammadiyah Sumatera Utara, saya bertanda tangan dibawah ini:

Nama : Dwi Nisyatul Wardah
NPM : 2209020175
Program Studi : Teknologi Informasi
Karya Ilmiah : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Muhammadiyah Sumatera Utara Hak Bedas Royalti Non-Eksekutif (*Non-Exclusive Royalty free Right*) atas penelitian skripsi saya yang berjudul:

**KLASIFIKASI MUSIM TANAM TANAMAN HORTIKULTURA
BERDASARKAN PARAMETER CUACA DAN TANAH
MENGUNAKAN ENSEMBLE LEARNING**

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Non-Eksekutif ini, Universitas Muhammadiyah Sumatera Utara berhak menyimpan, mengalih media, memformat, mengelola dalam bentuk database, merawat dan mempublikasikan Skripsi saya ini tanpa meminta izin dari saya selama tetap mencantumkan nama saya sebagai penulis dan sebagai pemegang dan atau sebagai pemilik hak cipta.

Demikian pernyataan ini dibuat dengan sebenarnya.

Medan, April 2026

Yang membuat pernyataan



Dwi Nisyatul Wardah

NPM. 2209020175

RIWAYAT HIDUP

DATA PRIBADI

Nama Lengkap : Dwi Nisyatul Wardah
Tempat dan Tanggal Lahir : Sungai Penuh, 27 Februari 2004
Alamat Rumah : JL. Yosudarso No.50A
Telepon/Faks/HP : 081378271861
E-mail : dwiwardah042@gmail.com.
Instansi Tempat Kerja :
Alamat Kantor :

DATA PENDIDIKAN

SD : SD 025/XI Desa Gedang TAMAT: 2016
SMP : SMPN 8 Kota Sungai Penuh TAMAT: 2019
SMA : SMAN 1 Kota Sungai Penuh TAMAT: 2022

KATA PENGANTAR



Assalamu'alaikum Warahmatullahi Wabarakatuh

Segala puji dan rasa syukur dipanjatkan kehadirat Allah SWT yang telah melimpahkan berkah serta karunia-Nya, sehingga penyelesaian naskah skripsi ini bisa terwujud dengan lancar. Selawat beserta salam senantiasa dihaturkan kepada junjungan Nabi Muhammad SAW, yang telah menuntun seluruh umat ke arah kebenaran serta menjadi sumber inspirasi dalam menempuh rangkaian riset keilmuan ini.

Penyusunan karya tulis ilmiah ini merupakan buah dari proses panjang yang dipenuhi beragam rintangan serta pengorbanan. Disadari sepenuhnya kalau tanpa dukungan dan tuntunan berbagai pihak, mustahil rasanya menuntaskan skripsi ini dengan baik. Karenanya, dengan segenap ketulusan hati, ucapan terima kasih disampaikan kepada:

1. Bapak Prof. Dr. Akrim, M.Pd., Rektor Universitas Muhammadiyah Sumatera Utara (UMSU)
2. Dr. Alkhowarizmi S.Kom., M.Kom selaku dekan Fakultas Ilmu Komputer dan Teknologi Informasi. Dr. Firahmi Rizky, S.Kom.M.Kom selaku Wakil Dekan I, Mhd. Basri, S.Si., M.Kom selaku Wakil Dekan III.
3. Fatma Sari Hutagalung, S.Kom., M.Kom, Selaku Ketua Prodi Teknologi Informasi Fakultas Ilmu Komputer dan Teknologi Informasi.
4. Bapak Okvi Nugroho, S.Kom., M.Kom Sekretaris Program Studi Teknologi Informasi.

5. Dr. Firahmi Rizky, S.Kom., M.Kom, selaku dosen pembimbing yang telah memberi arahan, masukan, dukungan serta bimbingan yang sangat berharga dalam penyusunan skripsi ini.
6. Amrullah, S.Kom., M.Kom, selaku dosen pembimbing akademik yang telah memberikan dukungan dan arahan kepada penulis dalam menempuh pendidikan.
7. Papa dan Mama tercinta penulis, yang senantiasa menjadi rumah paling teduh untuk pulang. Dengan kasih sayang yang tak pernah putus, doa yang tak pernah henti, serta dukungan dan motivasi yang tulus, papa dan mama telah membesarkan dan menguatkan penulis dalam setiap langkah kehidupan. Atas cinta, pengorbanan, dan keikhlasan yang diberikan, penulis dapat menjadi setiap proses kehidupan hingga mencapai ini dan untuk seterusnya.
8. Saudara – saudara tercinta penulis, Indah Chairunnisa beserta Suami Moh. Supriyadi yang senantiasa menjadi penyemangat dalam setiap Langkah penulis. Dengan kasih sayang yang tulus, dukungan yang menguatkan, motivasi yang tak pernah surut, serta doa doa yang selalu menyertai, kehadiran mereka menjadi sumber kekuatan dan penghibur bagi penulis dalam menyelesaikan skripsi ini.
9. Sahabat dan kerabat seperjuangan, Muhammad Ariq Adrian, Dinda Anantya, Aida Fadhila, Farhan Fanalty, Eka Syarif Maulana. Terima kasih telah menjadi sahabat yang luar biasa, membantu penulis selama masa perkuliahan, dan saling menguatkan dalam penyusunan skripsi ini. Juga kepada seluruh keluarga besar D1 atas momen- momen indah yang telah kita lalui bersama.

10. Sahabat SMP Penulis, Rahmanda Artamevia, Alifa Rachma Fadila, Salsabila Vira, Laysha Indah Meilani, Dinda Azahra yang senantiasa mendengar keluh kesah penulis selama penulisan skripsi.
11. Sahabat SMA penulis, Nasywa Fadhilla Hasri, Aprilli Celi Denich, Citra Mayang Sari yang senantiasa mendengar keluh kesah penulis selama penulisan skripsi.
12. Kepada diri sendiri, Dwi Nisyatul Wardah, terima kasih karena telah memilih untuk tetap bertahan ketika segalanya terasa berat. Terima kasih karena tidak menyerah meskipun sering kali merasa sendirian, lelah, dan ragu terhadap kemampuan diri sendiri. Terima kasih atas setiap air mata yang jatuh dalam diam, setiap doa yang terucap dalam keheningan, dan setiap langkah kecil yang terus diupayakan meski hati ingin berhenti. Perjalanan ini tidak selalu mudah, namun keberanian untuk terus melangkah adalah bukti bahwa diri ini kuat. Semoga semua proses yang telah dilalui menjadi pengingat bahwa setiap perjuangan tidak pernah sia-sia, dan bahwa diri ini pantas untuk bangga atas apa yang telah dicapai.

KLASIFIKASI MUSIM TANAM TANAMAN HORTIKULTURA BERDASARKAN PARAMETER CUACA DAN TANAH MENGUNAKAN ENSEMBLE LEARNING

ABSTRAK

Sektor pertanian memiliki peran strategis dalam mendukung ketahanan pangan dan stabilitas ekonomi di Indonesia. Salah satu faktor penting dalam keberhasilan produksi pertanian adalah ketepatan dalam menentukan musim tanam yang sesuai dengan kondisi lingkungan. Namun, perubahan iklim dan variabilitas parameter cuaca menyebabkan ketidakpastian dalam penentuan waktu tanam secara konvensional. Oleh karena itu, diperlukan pendekatan berbasis data untuk membantu penyuluh pertanian dalam memberikan rekomendasi musim tanam yang lebih akurat dan sistematis kepada petani. Penelitian ini bertujuan untuk mengembangkan sistem rekomendasi musim tanam tanaman hortikultura berdasarkan parameter cuaca dan tanah menggunakan metode machine learning Random Forest. Dataset yang digunakan diperoleh dari platform Kaggle yang mencakup parameter lingkungan seperti suhu, curah hujan, kelembapan, pH tanah, intensitas cahaya, serta kandungan unsur hara. Tahapan penelitian meliputi pra-pemrosesan data, transformasi label musim, pembagian data latih dan data uji, serta pelatihan model menggunakan algoritma Random Forest. Sistem kemudian diimplementasikan dalam bentuk aplikasi berbasis web yang digunakan oleh penyuluh pertanian sebagai alat bantu dalam proses analisis dan pemberian rekomendasi musim tanam. Hasil evaluasi model menunjukkan bahwa Random Forest mampu mencapai akurasi sebesar **61,3%**, dengan nilai precision sebesar **59,9%**, recall sebesar **61,3%**, dan F1-score sebesar **60,3%** pada data uji. Hasil tersebut menunjukkan bahwa model memiliki kemampuan yang cukup baik dalam mengklasifikasikan musim tanam berdasarkan parameter lingkungan yang kompleks. Dengan demikian, sistem yang dikembangkan dapat membantu penyuluh pertanian dalam memberikan rekomendasi musim tanam yang lebih objektif dan berbasis data, sehingga diharapkan dapat meningkatkan efektivitas penentuan waktu tanam dan mendukung produktivitas tanaman hortikultura.

Kata Kunci: machine learning, musim tanam, Random Forest, sistem rekomendasi, tanaman hortikultura.

CLASSIFICATION OF HORTICULTURAL PLANTING SEASONS BASED ON WEATHER AND SOIL PARAMETERS USING ENSEMBLE LEARNING

ABSTRACT

The agricultural sector plays a strategic role in supporting food security and economic stability in Indonesia. One of the key factors in agricultural productivity is the accuracy in determining the appropriate planting season based on environmental conditions. However, climate change and variability in weather parameters create uncertainty in determining optimal planting time using conventional approaches. Therefore, a data-driven approach is required to assist agricultural extension workers in providing more accurate and systematic planting season recommendations to farmers. This study aims to develop a horticultural planting season recommendation system based on weather and soil parameters using the Random Forest machine learning method.

The dataset used in this study was obtained from Kaggle, consisting of environmental parameters such as temperature, rainfall, humidity, soil pH, light intensity, and soil nutrient content. The research stages include data preprocessing, seasonal label transformation, data splitting into training and testing sets, and model training using the Random Forest algorithm. The system is implemented as a web-based application designed to assist agricultural extension workers in analyzing data and generating planting season recommendations.

The evaluation results show that the Random Forest model achieves an accuracy of **61.3%**, with a precision of **59.9%**, recall of **61.3%**, and F1-score of **60.3%** on the testing data. These results indicate that the model has a reasonable capability in classifying planting seasons based on complex environmental parameters.

In conclusion, the developed system can support agricultural extension workers in providing more objective and data-driven planting season recommendations, which is expected to improve planting time management and enhance horticultural crop productivity.

Keywords: machine learning, planting season, Random Forest, recommendation system, horticultural crops

DAFTAR ISI

ABSTRAK	i
ABSTRACT	ii
DAFTAR ISI	iii
DAFTAR TABEL	v
DAFTAR GAMBAR	vi
BAB I. PENDAHULUAN	
1.1 LATAR BELAKANG MASALAH	1
1.2 RUMUSAN MASALAH.....	2
1.3 BATASAN MASALAH.....	3
1.4 TUJUAN PENELITIAN.....	4
1.5 MANFAAT PENELITIAN.....	4
BAB II. LANDASAN TEORI	
2.1 Teknologi Informasi dalam Pengambilan Keputusan	6
2.2 Sistem Rekomendasi	6
2.3 Machine Learning.....	7
2.4 Ensemble Learning dan Algoritma Random Forest.....	8
2.5 Studi Literatur dan Kajian Literatur	9
2.6 Parameter Cuaca dan Tanah sebagai Data Sistem	11
2.7 Tanaman Hortikultura dan Musim Tanam.....	13
BAB III. METODOLOGI PENELITIAN	
3.1 Jenis dan Pendekatan Penelitian	15
3.2 Objek dan Ruang Lingkup Penelitian	16
3.3 Dataset dan Sumber Data.....	17
3.3.1 Deskripsi Dataset	17
3.3.2 Variabel Penelitian	19
3.3.3 Peran Dataset dalam Penelitian	20
3.4 Tahapan Penelitian	20
3.4.1 Alur Penelitian.....	20
3.4.2 Flowchart Penelitian	22
3.5 Pra-Pemrosesan Data.....	22
3.5.1 Pembersihan Data.....	23
3.5.2 Transformasi dan Encoding Data	23
3.5.3 Normalisasi Data	23
3.5.4 Transformasi Label Musim	24
3.5.5 Pembagian Data.....	25
3.6 Metode Random Forest	25
3.6.1 Konsep Dasar Random Forest.....	25
3.6.2 Kriteria Pemidahan Node (Gini Impurity)	26
3.6.3 Parameter Random Forest	26
3.6.4 Proses Pelatihan dan Prediksi Model.....	27
3.6.5 Feature Importance	27
3.6.6 Flowchart Random Forest	27
3.7 Perancangan Sistem Rekomendasi.....	28
3.7.1 Arsitektur Sistem.....	29
3.7.2 Alur Kerja Sistem Rekomendasi	30
3.7.3 Keluaran Sistem.....	30
3.7.4 Use Case Diagram	30
3.7.5 Activity Diagram	31

3.7.6 Class Diagram	33
3.7.7 Deployment Diagram.....	34
3.7.8 Perancangan Wireframe Antarmuka system.....	35
3.8 Metode Evaluasi Model.....	38
3.8.1 Confusion Matrix.....	38
3.8.2 Accuracy	39
3.8.3 Precision.....	39
3.8.4 Recall	39
3.8.5 F1-Score.....	39
3.9 Jadwal Penelitian	40
BAB IV. HASIL DAN PEMBAHASAN	
4.1 Implementasi Antarmuka.....	41
4.1.1 Implementasi Halaman Dashboard	41
4.1.2 Implementasi Halaman Musim Tanam	43
4.1.3 Implementasi Hasil Prediksi dan Rekomendasi.....	45
4.1.4 Implementasi Riwayat Analisis	47
4.1.5 Implementasi Management Data	49
4.1.6 Implementasi Halaman Data Holtikultura.....	50
4.1.7 Implementasi Pusat Pelatihan Model	52
4.2 Implementasi Sistem dan Arsitektur Integrasi Teknologi	53
4.3 Analisis Hasil Pra-pemrosesan Data dan Karakteristik Dataset	56
4.4 Konfigurasi dan Hasil Pelatihan Model Random Forest	58
4.5 Evaluasi Performa Model Menggunakan Metrik Klasifikasi	60
4.6 Analisis Feature Importance dan Kontribusi Parameter Lingkungan ..	62
4.7 Pengujian Sistem Rekomendasi Berbasis Skenario Nyat.....	64
4.8 Analisis Kestabilan, Keterbatasan, dan Potensi Pengembangan.....	65
BAB V. PENUTUP	
5.1 Kesimpulan	67
5.2 Saran	68
DAFTAR PUSTAKA	70

DAFTAR TABEL

Tabel 2.1 Penelitian Terdahulu.....	10
Tabel 3.1 Variabel Dataset Penelitian.....	18
Tabel 3.2 Jadwal Penelitian.....	40
Tabel 4.1 Komponen Sistem	55
Tabel 4.2 Parameter Tanah/Cuasa	57
Tabel 4.3 Hyperparameter Model.....	59
Tabel 4.4 Peringkat Fitur	63
Tabel 4.5 Input Skenario	65

DAFTAR GAMBAR

Gambar 3.1 Flowchart Tahapan Penelitian	22
Gambar 3.2 Flowchart Random Forest	28
Gambar 3.3 Use Case Diagram Sistem Rekomendasi Musim Tanam	31
Gambar 3.4 Activity Diagram Sistem Rekomendasi Musim Tanam	32
Gambar 3.5 Class Diagram Sistem Rekomendasi Musim Tanam.....	33
Gambar 3.6 Deployment Diagram Sistem Rekomendasi Musim Tanam	34
Gambar 3.7 Wireframe Halaman Dashboard	36
Gambar 3.8 Wireframe Halaman Analisis dan Rekomendasi.....	36
Gambar 3.9 Wireframe Halaman Hasil Analisis	36
Gambar 3.10 Wireframe Halaman Riwayat Analisis	37
Gambar 3.11 Wireframe Halaman Management Data	37
Gambar 3.12 Wireframe Halaman Data Holtikultura.....	37
Gambar 3.13 Wireframe Halaman Pusat Pelatihan Model	38
Gambar 4.1 Halaman Dashboard.....	41
Gambar 4.2 Halaman Analisis Lahan & Rekomendasi	44
Gambar 4.3 Hasil Prediksi dan Rekomendasi	45
Gambar 4.4 Halaman Riwayat Analisis	47
Gambar 4.5 Halaman Management Data	49
Gambar 4.6 Halaman Data Holtikultura	50
Gambar 4.7 Halaman Pelatihan Model	52

BAB I

PENDAHULUAN

1.1. Latar Belakang Masalah

Sektor pertanian memiliki peran strategis dalam menjaga ketahanan pangan dan stabilitas ekonomi, khususnya di negara berkembang seperti Indonesia. Keberhasilan produksi pertanian sangat dipengaruhi oleh ketepatan dalam menentukan waktu tanam dan pemilihan musim yang sesuai dengan kondisi lingkungan. Namun, perubahan iklim, variabilitas curah hujan, serta ketidakstabilan suhu menyebabkan pola musim menjadi semakin sulit diprediksi secara konvensional (Shastri et al., 2025). Kondisi ini menuntut pendekatan yang lebih berbasis data dalam pengambilan keputusan musim tanam.

Dalam praktiknya, banyak petani masih mengandalkan pengalaman empiris atau kebiasaan turun-temurun dalam menentukan waktu tanam. Pendekatan tersebut tidak selalu mampu mengakomodasi perubahan lingkungan yang dinamis dan kompleks (Prity et al., 2024). Oleh karena itu, diperlukan peran pihak yang memiliki kapasitas analisis dan akses terhadap informasi teknis yang lebih komprehensif, yaitu penyuluh pertanian. Penyuluh berfungsi sebagai penghubung antara perkembangan ilmu pengetahuan dengan praktik di lapangan, serta berperan dalam memberikan rekomendasi teknis kepada petani (Yanfika et al., 2024).

Seiring berkembangnya teknologi pengolahan data, metode machine learning telah banyak digunakan dalam bidang pertanian untuk melakukan klasifikasi dan prediksi berbasis parameter lingkungan seperti suhu, curah hujan, kelembapan, pH tanah, serta kandungan unsur hara (Shastri et al., 2025). Model berbasis ensemble

seperti Random Forest terbukti mampu menangani data multivariat dan memberikan tingkat akurasi yang baik dalam berbagai studi pertanian modern (Prity et al., 2024). Penggunaan pendekatan ini memungkinkan analisis hubungan kompleks antar variabel lingkungan yang sulit ditangani dengan metode statistik sederhana.

Selain aspek akurasi, sistem rekomendasi juga perlu menyediakan penjelasan terhadap hasil prediksi agar keputusan yang diambil dapat dipahami secara rasional oleh pengguna profesional. Transparansi model menjadi penting agar penyuluh dapat menjelaskan dasar rekomendasi musim tanam kepada petani dengan argumentasi yang jelas dan terukur (Hussein et al., 2024). Dengan demikian, sistem tidak hanya berfungsi sebagai alat prediksi, tetapi juga sebagai alat bantu keputusan yang mendukung proses komunikasi teknis di lapangan.

Berdasarkan uraian tersebut, penelitian ini mengembangkan sistem rekomendasi musim tanam berbasis machine learning yang ditujukan untuk penyuluh pertanian sebagai pengguna utama. Sistem ini dirancang untuk mengolah parameter cuaca dan kondisi tanah guna menghasilkan rekomendasi musim tanam yang lebih akurat dan berbasis data, sehingga dapat membantu penyuluh dalam memberikan arahan yang lebih tepat kepada petani.

1.2. Rumusan Masalah

Berdasarkan latar belakang tersebut, rumusan masalah dalam penelitian ini adalah:

1. Bagaimana penerapan metode *Random Forest* dalam membangun sistem rekomendasi musim tanam tanaman hortikultura berdasarkan parameter cuaca dan tanah?

2. Seberapa menguji performa model *Random Forest* dalam memberikan rekomendasi musim tanam untuk berbagai jenis tanaman hortikultura?
3. Bagaimana merancang dan membangun sistem rekomendasi musim tanam tanaman hortikultura berbasis sistem informasi yang dapat digunakan oleh penyuluh pertanian dalam memberikan rekomendasi kepada petani?

1.3. Batasan Masalah

Agar penelitian lebih terarah, batasan masalah ditetapkan sebagai berikut:

1. Tanaman hortikultura terbatas pada yang terdapat dalam dataset, seperti stroberi, semangka, anggur, arugula, bit, bayam, kale, selada, radicchio, tomat, terong, asparagus, cabai, kol, mentimun, kentang, kembang kol, brokoli, dan kacang polong.
2. Parameter yang digunakan meliputi: fertility, photoperiod, temperature, rainfall, pH, Light_Hours, light_intensity, Rh, nitrogen, phosphorus, potassium, yield, category pH, soil type, season, serta rasio nutrisi (N_Ratio, P_Ratio, K_Ratio).
3. Penelitian hanya menggunakan metode *ensemble learning* Random Forest sebagai model utama.
4. Lingkup penelitian difokuskan pada kondisi agroklimat Indonesia dan tidak mencakup pengujian lapangan berskala luas.
5. Sistem yang dibangun berupa prototipe dan tidak mencakup integrasi penuh dengan platform eksternal.

1.4. Tujuan Penelitian

Berdasarkan rumusan masalah yang telah dirumuskan, maka tujuan dari penelitian ini adalah sebagai berikut:

1. Menerapkan metode *Random Forest* dengan memanfaatkan data parameter cuaca dan tanah sebagai dasar dalam membangun sistem rekomendasi musim tanam tanaman hortikultura.
2. Mengetahui performa model *Random Forest* dalam memberikan rekomendasi musim tanam tanaman hortikultura untuk berbagai jenis tanaman berdasarkan hasil pengujian model.
3. Merancang dan membangun sistem rekomendasi musim tanam tanaman hortikultura berbasis sistem informasi yang mengintegrasikan pengolahan data dan model *Random Forest* sebagai alat bantu keputusan bagi penyuluh pertanian dalam memberikan rekomendasi musim tanam kepada petani.

1.5. Manfaat Penelitian

Berdasarkan penelitian yang dilakukan, manfaat yang diharapkan dari penelitian ini adalah sebagai berikut:

1. Menambah literatur ilmiah mengenai penerapan *machine learning*, khususnya *Random Forest*, untuk rekomendasi musim tanam tanaman hortikultura.
2. Memberikan kontribusi metodologis dalam pengolahan data pertanian yang kompleks menggunakan *ensemble learning*.
3. Membantu penyuluh pertanian dalam memberikan rekomendasi musim tanam yang lebih akurat dan berbasis data kepada petani, sehingga diharapkan dapat meningkatkan produktivitas dan efisiensi penggunaan sumber daya pertanian.

4. Menjadi dasar pengembangan sistem pertanian digital yang dapat diintegrasikan dengan aplikasi rekomendasi pertanian nasional.
5. Mendukung strategi ketahanan pangan Indonesia melalui pemanfaatan teknologi berbasis data dalam sektor hortikultura.

BAB II

LANDASAN TEORI

2.1 Teknologi Informasi dalam Pengambilan Keputusan

Teknologi informasi merupakan bidang yang berkaitan dengan pemanfaatan sistem komputer untuk mengelola, memproses, dan menyajikan data menjadi informasi yang berguna. Dalam konteks pengambilan keputusan, teknologi informasi berperan sebagai alat pendukung yang mampu membantu pengguna dalam menganalisis data secara objektif dan sistematis. Pemanfaatan teknologi informasi memungkinkan proses pengambilan keputusan dilakukan secara lebih cepat dan akurat dibandingkan pendekatan konvensional yang mengandalkan intuisi atau pengalaman semata (Sarker, 2021).

Perkembangan teknologi informasi mendorong lahirnya sistem berbasis data (*data-driven system*), di mana keputusan dihasilkan berdasarkan analisis data historis dan kondisi aktual. Sistem semacam ini sangat relevan diterapkan pada bidang yang memiliki tingkat ketidakpastian tinggi, seperti pertanian, yang sangat dipengaruhi oleh faktor lingkungan dan cuaca (FAO, 2021). Oleh karena itu, teknologi informasi menjadi fondasi utama dalam pengembangan sistem pendukung keputusan dan sistem rekomendasi.

2.2 Sistem Rekomendasi

Sistem rekomendasi merupakan sistem berbasis komputer yang dirancang untuk memberikan saran atau rekomendasi kepada pengguna berdasarkan data, model, atau pengetahuan tertentu. Tujuan utama dari sistem rekomendasi adalah

membantu pengguna dalam mengambil keputusan dengan cara menyederhanakan proses analisis terhadap sejumlah alternatif yang tersedia (Ricci et al., 2022).

Dalam beberapa tahun terakhir, sistem rekomendasi tidak hanya digunakan pada bidang komersial seperti *e-commerce*, tetapi juga mulai diterapkan pada sektor lain, termasuk pertanian. Pada sektor pertanian, sistem rekomendasi dapat dimanfaatkan untuk membantu petani dalam menentukan keputusan strategis, seperti pemilihan jenis tanaman atau penentuan waktu tanam yang sesuai dengan kondisi lingkungan. Sistem rekomendasi berbasis data dinilai lebih adaptif karena mampu menyesuaikan rekomendasi dengan perubahan kondisi yang terjadi.

2.3 Machine Learning

Machine learning merupakan cabang dari kecerdasan buatan (*artificial intelligence*) yang memungkinkan sistem komputer mempelajari pola dari data dan membuat keputusan atau prediksi tanpa diprogram secara eksplisit. Dalam *supervised learning*, model dilatih menggunakan data yang memiliki label sebagai acuan untuk mempelajari hubungan antara variabel input dan output (Sarker, 2021).

Secara umum, proses *machine learning* bertujuan untuk meminimalkan kesalahan prediksi antara nilai aktual dan nilai hasil prediksi model. Kesalahan tersebut dapat direpresentasikan menggunakan fungsi kerugian (*loss function*), yang secara sederhana dirumuskan sebagai:

$$L = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (1)$$

Dimana :

- a. y_i adalah nilai *actual*
- b. \hat{y}_i adalah nilai prediksi model

c. n adalah jumlah data

Pendekatan *machine learning* sangat sesuai digunakan pada permasalahan pertanian karena mampu memodelkan hubungan yang kompleks antara parameter cuaca dan tanah dengan kondisi musim tanam (Khan et al., 2022).

2.4 Ensemble Learning dan Algoritma Random Forest

Ensemble learning merupakan pendekatan dalam *machine learning* yang mengombinasikan beberapa model prediktif untuk meningkatkan akurasi dan stabilitas hasil prediksi dibandingkan satu model tunggal. Prinsip utama *ensemble learning* adalah bahwa kombinasi beberapa *weak learners* dapat menghasilkan model yang lebih kuat (*strong learner*) (Zhou, 2021).

Salah satu algoritma *ensemble learning* yang paling banyak digunakan adalah *Random Forest*. *Random Forest* bekerja dengan membangun sejumlah *decision tree* pada subset data dan subset fitur yang dipilih secara acak (*bagging*). Hasil prediksi akhir diperoleh melalui mekanisme *majority voting* sebagai berikut:

$$\hat{y} = \text{mode} \{T_1(x), T_2(x), \dots, T_k(x)\} \quad (2)$$

Pemilihan pemisahan node pada setiap *decision tree* dilakukan menggunakan ukuran *impurity* seperti *Gini Index*, yang dirumuskan sebagai:

$$Gini = 1 - \sum_{i=1}^c p_i^2 \quad (3)$$

Dimana p_i adalah proporsi sampel pada kelas ke- i .

Random Forest memiliki keunggulan dalam menangani data berdimensi besar, tidak sensitif terhadap *outlier*, serta mampu memodelkan hubungan nonlinier. Oleh karena itu, algoritma ini banyak digunakan dalam prediksi dan pengambilan keputusan berbasis data pertanian (Sarker, 2021; Khan et al., 2022).

Selain itu, *Random Forest* juga mampu memberikan informasi mengenai *feature importance*, yang menunjukkan fitur mana yang paling berpengaruh pada proses prediksi. Nilai *importance* dihitung berdasarkan rata-rata penurunan impurity pada seluruh *tree*:

$$FI_j = \frac{1}{B} \sum_{b=1}^B \Delta I_b(j) \quad (4)$$

Random Forest banyak digunakan karena beberapa kelebihanannya, seperti mampu menangani data berdimensi besar, tidak sensitif terhadap outlier, dan memiliki akurasi tinggi pada data non-linear. Namun, model ini memiliki kekurangan seperti interpretasi yang lebih sulit dan waktu komputasi yang lebih tinggi pada dataset besar (Patel et al., 2023). Berdasarkan karakteristik tersebut, *Random Forest* sangat sesuai untuk penelitian ini karena data cuaca dan tanah memiliki pola non-linear dan interaksi variabel yang kompleks, sehingga memerlukan metode prediksi yang stabil dan robust.

2.5 Studi Terkait dan Kajian Literatur

Berbagai penelitian menunjukkan bahwa metode *ensemble learning* efektif digunakan dalam prediksi dan pengambilan keputusan di bidang pertanian. Sarker (2021) melaporkan bahwa algoritma *Random Forest* dan *XGBoost* mampu menangani pola data pertanian yang bersifat *non-linear*, khususnya pada parameter cuaca dan tanah, sehingga menghasilkan akurasi prediksi yang lebih tinggi dibandingkan model tunggal.

Penelitian oleh Khan et al. (2022) menunjukkan bahwa *Random Forest* mampu mencapai akurasi lebih dari 85% dalam memprediksi produktivitas tanaman hortikultura menggunakan variabel suhu, curah hujan, pH tanah, dan

nutrisi NPK. Penelitian lain oleh Li dan Chen (2021) juga menunjukkan bahwa metode *ensemble learning* memberikan hasil yang lebih akurat dalam penentuan waktu tanam optimal dibandingkan pendekatan *regresi* tradisional.

Dalam konteks Indonesia, Haryanto dan Sari (2022) membuktikan bahwa *Random Forest* lebih efektif dibandingkan *decision tree* dalam klasifikasi kesesuaian lahan pertanian berdasarkan parameter tanah. Selain itu, Patel et al. (2023) menyatakan bahwa integrasi parameter cuaca dan tanah dalam satu model prediksi mampu mengurangi kesalahan akibat variabilitas iklim.

Meskipun demikian, sebagian besar penelitian sebelumnya masih berfokus pada prediksi hasil panen atau kesesuaian lahan. Penelitian yang secara khusus mengembangkan sistem rekomendasi musim tanam tanaman hortikultura dengan mengintegrasikan parameter cuaca dan tanah secara bersamaan masih terbatas. Oleh karena itu, penelitian ini bertujuan mengisi celah tersebut dengan membangun sistem rekomendasi musim tanam hortikultura berbasis *Random Forest*.

Tabel 2.1. Penelitian Terdahulu

No	Penulis dan Tahun	Judul Penelitian	Metode Penelitian	Data yang Digunakan	Hasil Penelitian
1	Sarker (2021)	<i>Machine Learning Agriculture : A Review</i>	Random Forest, XGBoost	Data cuaca dan tanah	Ensemble Learning efektif menangani data pertanian non-linear
2	Khan et al. (2022)	Crop Yield Prediction Using Random Forest	Random Forest	Suhu, Curah hujan, pH,NPK	Akurasi prediksi >85%
3	Li & Chen	Crop Season Prediction Using ML	XGBoost	Suhu, Kelembapan, Curah Hujan	Akurasi meningkat dibanding regresi

4	Rahman et al. (2021)	Climate and Soil Based Crop Prediction	Random Forest	Data cuaca dan pH tanah	Integrasi cuaca-tanah meningkatkan performa
5	Zhang et al.	Horticultural Yield Prediction	Random Forest	Cahaya, suhu, tanah	Intensitas Cahaya dan pH paling berpengaruh
6	Haryanto dan sari (2022)	Kesesuaian Lahan Pertanian	Random Forest	Data tanah	RF lebih akurat dari decision tree

Berdasarkan penelitian-penelitian terdahulu tersebut, dapat disimpulkan bahwa algoritma *Random Forest* dan metode *ensemble learning* banyak digunakan dalam bidang pertanian karena mampu memberikan akurasi prediksi yang tinggi. Namun, sebagian besar penelitian masih berfokus pada prediksi hasil panen atau kesesuaian lahan, dan belum secara spesifik mengembangkan sistem rekomendasi musim tanam hortikultura yang mengintegrasikan parameter cuaca dan tanah secara bersamaan.

2.6 Parameter Cuaca dan Tanah sebagai Data Sistem

Parameter cuaca dan tanah merupakan faktor lingkungan utama yang memengaruhi pertumbuhan dan produktivitas tanaman hortikultura. Dalam konteks sistem berbasis teknologi informasi, parameter-parameter tersebut diperlakukan sebagai data input yang akan diolah oleh sistem untuk menghasilkan rekomendasi musim tanam. Pendekatan ini memungkinkan pengambilan keputusan dilakukan secara objektif berdasarkan data, bukan hanya berdasarkan pengalaman atau perkiraan subjektif.

Parameter cuaca yang umum digunakan dalam analisis pertanian meliputi suhu udara, curah hujan, dan kelembapan udara. Suhu udara berpengaruh terhadap proses fisiologis tanaman, seperti fotosintesis dan respirasi. Curah hujan berkaitan

dengan ketersediaan air bagi tanaman, sementara kelembapan udara memengaruhi tingkat penguapan dan risiko serangan penyakit. Variasi pada parameter cuaca tersebut dapat menyebabkan perbedaan kondisi tumbuh tanaman pada setiap periode waktu (FAO, 2021).

Selain parameter cuaca, kondisi tanah juga memiliki peran penting dalam menentukan keberhasilan budidaya tanaman hortikultura. Parameter tanah yang sering digunakan sebagai data antara lain pH tanah dan ketersediaan unsur hara. pH tanah memengaruhi kemampuan tanaman dalam menyerap nutrisi, sedangkan unsur hara berperan langsung dalam mendukung pertumbuhan vegetatif dan generatif tanaman. Tanah dengan kondisi yang tidak sesuai dapat menghambat pertumbuhan tanaman meskipun faktor cuaca berada pada kondisi optimal.

Dalam sistem rekomendasi berbasis *machine learning*, parameter cuaca dan tanah direpresentasikan dalam bentuk data numerik yang dapat diproses oleh model. Setiap parameter berfungsi sebagai fitur (*feature*) yang memberikan informasi tertentu kepada sistem. Kombinasi beberapa parameter tersebut memungkinkan model untuk mempelajari pola hubungan antara kondisi lingkungan dan musim tanam yang sesuai. Dengan demikian, sistem dapat menghasilkan rekomendasi musim tanam yang lebih akurat dan adaptif terhadap perubahan kondisi lingkungan.

Pendekatan berbasis data ini dinilai lebih efektif dibandingkan metode konvensional karena mampu mengintegrasikan berbagai parameter lingkungan secara bersamaan. Oleh karena itu, pemilihan parameter cuaca dan tanah sebagai data input merupakan komponen penting dalam pengembangan sistem rekomendasi musim tanam tanaman hortikultura.

2.7 Tanaman Hortikultura dan Musim Tanam

Tanaman hortikultura merupakan kelompok tanaman yang mencakup sayuran, buah-buahan, dan tanaman hias yang memiliki nilai ekonomi dan gizi yang tinggi. Tanaman hortikultura umumnya memiliki karakteristik sensitif terhadap perubahan kondisi lingkungan, sehingga membutuhkan pengelolaan budidaya yang lebih tepat dibandingkan tanaman pangan lainnya (FAO, 2022).

Salah satu aspek penting dalam budidaya tanaman hortikultura adalah penentuan musim tanam. Musim tanam merupakan periode waktu yang dianggap paling sesuai untuk memulai penanaman suatu komoditas berdasarkan kondisi lingkungan yang mendukung pertumbuhan tanaman. Penentuan musim tanam yang tepat dapat meningkatkan peluang keberhasilan budidaya, meningkatkan hasil panen, serta mengurangi risiko kegagalan akibat kondisi cuaca yang tidak sesuai.

Perubahan iklim menyebabkan pola musim tanam menjadi semakin sulit diprediksi. Ketidakpastian curah hujan, fluktuasi suhu, dan perubahan pola musim dapat menyebabkan ketidaksesuaian antara waktu tanam dan kebutuhan tanaman. Kondisi ini berdampak pada menurunnya produktivitas tanaman hortikultura apabila penanaman dilakukan pada waktu yang tidak tepat (FAO, 2021).

Dalam konteks penelitian ini, tanaman hortikultura diposisikan sebagai objek dari sistem rekomendasi. Sistem tidak berfokus pada teknik budidaya secara agronomis, tetapi pada penentuan waktu tanam yang sesuai berdasarkan analisis data parameter cuaca dan tanah. Dengan memanfaatkan sistem rekomendasi berbasis *machine learning*, informasi mengenai karakteristik tanaman dan kondisi lingkungan dapat diolah untuk menghasilkan rekomendasi musim tanam yang lebih presisi.

Pengembangan sistem rekomendasi musim tanam diharapkan dapat membantu petani dalam menentukan waktu tanam yang optimal untuk berbagai jenis tanaman hortikultura. Dengan demikian, keputusan yang diambil tidak hanya berdasarkan kebiasaan atau pengalaman, tetapi didukung oleh analisis data yang sistematis dan terukur.

BAB III

METODOLOGI PENELITIAN

3.1 Jenis dan Pendekatan Penelitian

Penelitian ini merupakan penelitian eksperimental, yaitu penelitian yang dilakukan dengan menguji kinerja suatu metode untuk memperoleh hasil yang optimal. Dalam penelitian ini, eksperimen dilakukan dengan membangun dan menguji *model machine learning* menggunakan algoritma *Random Forest* untuk menghasilkan sistem rekomendasi musim tanam tanaman hortikultura. Proses penelitian meliputi tahap pelatihan model, pengujian model, serta evaluasi kinerja model berdasarkan hasil pengujian.

Pendekatan yang digunakan dalam penelitian ini adalah pendekatan kuantitatif, karena penelitian ini mengolah dan menganalisis data dalam bentuk numerik. Data yang digunakan berupa parameter cuaca dan tanah, seperti suhu udara, curah hujan, pH tanah, kelembapan, serta kandungan unsur hara. Hasil penelitian juga dinyatakan dalam bentuk nilai numerik, sehingga pendekatan kuantitatif dinilai sesuai untuk mendukung tujuan penelitian.

Metode penelitian yang digunakan adalah *machine learning* dengan pendekatan *supervised learning*, karena model dilatih menggunakan data yang telah memiliki label atau target *output*. Algoritma yang digunakan adalah *Random Forest*, yang bekerja dengan menggabungkan banyak *decision tree* untuk menghasilkan prediksi akhir. *Random Forest* dipilih karena mampu menangani data dengan pola non-linear, mengurangi risiko *overfitting*, serta memberikan performa yang baik pada data cuaca dan tanah yang kompleks.

3.2 Objek dan Ruang Lingkup Penelitian

Objek dalam penelitian ini adalah tanaman hortikultura beserta musim tanamnya yang dipengaruhi oleh kondisi lingkungan. Tanaman hortikultura yang menjadi objek penelitian meliputi berbagai jenis tanaman buah dan sayuran, seperti stroberi, semangka, anggur, tomat, cabai, mentimun, kentang, brokoli, dan sayuran daun lainnya sesuai dengan dataset yang digunakan. Selain tanaman, objek penelitian juga mencakup parameter cuaca dan tanah yang berperan dalam menentukan kesesuaian musim tanam, seperti suhu udara, curah hujan, kelembapan, intensitas cahaya, pH tanah, tingkat kesuburan tanah, serta kandungan unsur *hara nitrogen*, *fosfor*, dan *kalium*.

Penelitian ini berfokus pada pemanfaatan data historis parameter lingkungan tersebut untuk menghasilkan rekomendasi musim tanam yang optimal melalui pendekatan *machine learning*. Dengan demikian, objek penelitian tidak hanya terbatas pada komoditas tanaman, tetapi juga mencakup hubungan antara karakteristik lingkungan dan waktu tanam yang sesuai.

Ruang lingkup penelitian ini dibatasi untuk memastikan fokus penelitian tetap terarah dan sesuai dengan tujuan yang ingin dicapai. Penelitian hanya menggunakan dataset sekunder yang berisi data cuaca, tanah, dan tanaman hortikultura tanpa melakukan pengambilan data langsung di lapangan. Model yang digunakan dalam penelitian ini dibatasi pada algoritma *Random Forest* sebagai metode utama dalam membangun sistem rekomendasi musim tanam.

Parameter yang dianalisis terbatas pada variabel yang tersedia dalam dataset, meliputi parameter cuaca (suhu, curah hujan, kelembapan, lama penyinaran, dan intensitas cahaya) serta parameter tanah (pH, jenis tanah, tingkat kesuburan, dan

kandungan NPK). Output yang dihasilkan dari penelitian ini berupa rekomendasi musim tanam tanaman hortikultura berdasarkan hasil prediksi model.

Penelitian ini tidak membahas aspek ekonomi, kebijakan pertanian, maupun pengujian sistem secara langsung di lapangan. Evaluasi penelitian difokuskan pada kinerja model berdasarkan hasil pengujian data dan metrik evaluasi yang telah ditentukan. Dengan ruang lingkup tersebut, penelitian diharapkan mampu memberikan gambaran yang jelas mengenai penerapan *Random Forest* dalam sistem rekomendasi musim tanam berbasis data.

3.3 Dataset dan Sumber Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari platform *Kaggle*, yaitu platform berbagi dataset yang banyak digunakan dalam penelitian dan pengembangan *machine learning*. Dataset *Kaggle* dipilih karena mudah diakses, terstruktur dengan baik, serta sering digunakan sebagai acuan dalam penelitian berbasis data. Data yang digunakan berisi informasi historis terkait tanaman hortikultura, parameter cuaca, dan parameter tanah yang relevan untuk analisis musim tanam.

Penggunaan dataset dari *Kaggle* memungkinkan penelitian dilakukan secara efisien tanpa pengambilan data langsung di lapangan, serta tetap memberikan representasi kondisi lingkungan yang dibutuhkan untuk membangun model prediksi dan rekomendasi musim tanam.

3.3.1 Deskripsi Dataset

Dataset yang digunakan dalam penelitian ini diperoleh dari platform *Kaggle* dengan nama dataset Soil Nutrients Dataset. Dataset ini berisi data lingkungan yang

berkaitan dengan pertumbuhan tanaman hortikultura berdasarkan parameter cuaca dan kondisi tanah.

Dataset terdiri dari 15.400 data observasi dengan 19 fitur yang merepresentasikan berbagai parameter lingkungan yang memengaruhi musim tanam tanaman hortikultura. Fitur tersebut mencakup parameter cuaca seperti suhu (Temperature), curah hujan (Rainfall), kelembapan udara (Rh), lama penyinaran matahari (Light_Hours), serta intensitas cahaya (Light_Intensity). Selain itu, dataset juga memuat parameter tanah seperti pH tanah, jenis tanah (Soil_Type), tingkat kesuburan tanah (Fertility), serta kandungan unsur hara nitrogen (Nitrogen), fosfor (Phosphorus), dan kalium (Potassium).

Dataset juga mencakup informasi tambahan seperti rasio unsur hara (N_Ratio, P_Ratio, K_Ratio), kategori pH tanah (Category_pH), serta data hasil panen (Yield) sebagai informasi pendukung. Variabel target dalam penelitian ini adalah Season, yang merepresentasikan kategori musim tanam tanaman hortikultura. Variabel ini digunakan sebagai label klasifikasi dalam model machine learning Random Forest untuk menentukan musim tanam yang sesuai berdasarkan kondisi parameter cuaca dan tanah yang diberikan.

Tabel 3.1 Variabel Dataset Penelitian

No	Fitur	Keterangan
1	Name	Nama tanaman hortikultura
2	Fertility	Tingkat kesuburan tanah
3	Photoperiod	Periode penyinaran tanaman
4	Temperature	Suhu udara
5	Rainfall	Curah hujan
6	pH	Tingkat keasaman tanah
7	Light_Hours	Durasi penyinaran

8	Light_Intensity	Intensitas cahaya
9	Rh	Kelembapan udara
10	Nitrogen	Kandungan nitrogen
11	Phosphorus	Kandungan fosfor
12	Potassium	Kandungan kalium
13	Yield	Hasil panen
14	Category_pH	Kategori pH tanah
15	Soil_Type	Jenis tanah
16	Season	Musim tanam (target klasifikasi)
17	N_Ratio	Rasio nitrogen
18	P_Ratio	Rasio fosfor
19	K_Ratio	Rasio kalium

3.3.2 Variabel Penelitian

Variabel yang digunakan dalam penelitian ini dibagi menjadi dua kelompok, yaitu *variable input* dan *variable output*.

1. Variabel Input

Variabel *input* merupakan parameter lingkungan yang memengaruhi musim tanam tanaman hortikultura, meliputi :

- a. Parameter cuaca: Suhu udara, curah hujan, kelembapan relative, lama penyinaran (photoperiod), intensitas Cahaya, dan jumlah jam Cahaya.
- b. Parameter tanah: pH tanah, kategori pH, jenis tanah, Tingkat kesuburan tanah, kandungan nitrogen, fosfor, dan kalium, serta rasio unsur hara (N_ratio,P_ratio,K_ratio).

2. Variabel Output

Variabel output dalam penelitian ini adalah musim tanam tanaman hortikultura, yang menjadi dasar dalam sistem rekomendasi. Selain itu, dataset juga

memuat informasi hasil panen (*yield*) yang dapat digunakan sebagai data pendukung dalam analisis performa tanaman.

3.3.3 Peran Dataset dalam Penelitian

Dataset ini berperan sebagai dasar utama dalam seluruh proses penelitian, mulai dari tahap *preprocessing*, pelatihan model *Random Forest*, hingga evaluasi hasil prediksi. Dengan memanfaatkan dataset yang memuat berbagai parameter cuaca dan tanah, penelitian ini dapat menganalisis pola hubungan antara kondisi lingkungan dan musim tanam tanaman hortikultura secara komprehensif. Hasil dari pemodelan tersebut kemudian digunakan untuk membangun sistem rekomendasi musim tanam yang bersifat *data-driven* dan mudah diaplikasikan.

3.4 Tahapan Penelitian

Tahapan penelitian merupakan langkah-langkah sistematis yang dilakukan untuk mencapai tujuan penelitian. Pada penelitian ini, tahapan dimulai dari pengumpulan dataset hingga menghasilkan rekomendasi musim tanam tanaman hortikultura berbasis *Random Forest*. Tahapan penelitian dirancang agar proses pengolahan data dan pemodelan dapat berjalan secara terstruktur dan mudah dipahami.

3.4.1 Alur Penelitian

Alur penelitian pada sistem rekomendasi musim tanam tanaman hortikultura ini terdiri dari beberapa tahapan utama, yaitu:

a. Pengumpulan Data

Tahap awal penelitian adalah pengumpulan dataset yang diperoleh dari Kaggle. Dataset ini berisi data tanaman hortikultura, parameter cuaca, dan parameter tanah yang digunakan sebagai dasar analisis.

b. Pra-pemrosesan Data

Data yang telah dikumpulkan kemudian diproses untuk memastikan kualitas data. Tahap ini meliputi pembersihan data, penanganan nilai kosong (*missing value*), transformasi data, serta encoding data kategorikal agar dapat diproses oleh algoritma *machine learning*.

c. Pembagian Data

Dataset dibagi menjadi dua bagian, yaitu data latih (*training data*) dan data uji (*testing data*). Data latih digunakan untuk membangun model *Random Forest*, sedangkan data uji digunakan untuk mengevaluasi kinerja model.

d. Pelatihan Model Random Forest

Pada tahap ini, model *Random Forest* dilatih menggunakan data latih. Proses pelatihan bertujuan untuk mempelajari pola hubungan antara parameter cuaca dan tanah dengan musim tanam tanaman hortikultura.

e. Pengujian Model

Model yang telah dilatih kemudian diuji menggunakan data uji untuk mengetahui kemampuan model dalam memberikan prediksi musim tanam secara akurat.

f. Pengujian Model

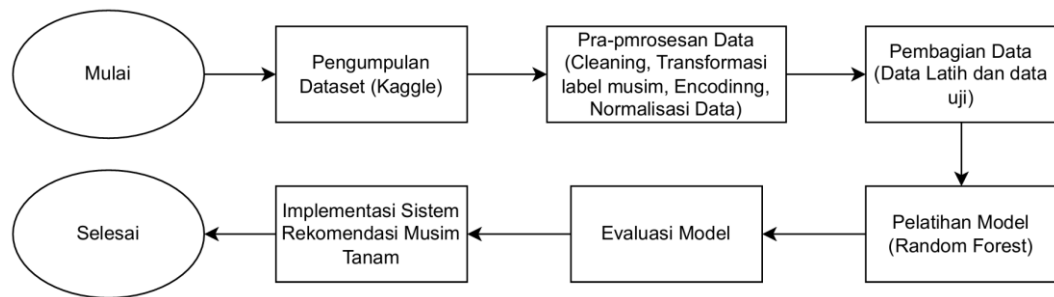
Model yang telah dilatih kemudian diuji menggunakan data uji untuk mengetahui kemampuan model dalam memberikan prediksi musim tanam secara akurat.

g. System Rekomendasi Musim Tanam

Tahap akhir adalah penyajian hasil prediksi dalam bentuk rekomendasi musim tanam tanaman hortikultura berdasarkan parameter cuaca dan tanah yang diberikan.

3.4.2 Flowchart Penelitian

Berikut adalah *flowchart* tahapan penelitian yang menggambarkan alur kerja sistem rekomendasi musim tanam tanaman hortikultura:



Gambar 3.1 *Flowchart* Tahapan Penelitian

Flowchart tersebut menunjukkan bahwa proses penelitian dilakukan secara berurutan dan terintegrasi, mulai dari pengolahan data hingga menghasilkan rekomendasi musim tanam. Dengan alur ini, sistem yang dibangun diharapkan mampu memberikan rekomendasi yang akurat dan mudah dipahami oleh pengguna.

3.5 Pra-Pemrosesan Data

Pra-pemrosesan data merupakan tahap penting dalam penelitian ini untuk memastikan kualitas data sebelum digunakan dalam proses pelatihan model. Tahap ini bertujuan untuk mengurangi kesalahan, meningkatkan konsistensi data, serta menyesuaikan format data agar dapat diproses dengan baik oleh algoritma *Random Forest*. Pra-pemrosesan data dilakukan berdasarkan dataset yang diperoleh dari *Kaggle* dan disesuaikan dengan kebutuhan penelitian.

3.5.1 Pembersihan Data (*Data Cleaning*)

Tahap pembersihan data dilakukan untuk mengatasi permasalahan pada dataset, seperti data kosong (*missing value*), data duplikat, dan nilai yang tidak wajar. Pada tahap ini, data yang tidak lengkap diperiksa dan ditangani dengan metode yang sesuai, seperti penghapusan data atau pengisian nilai berdasarkan karakteristik data yang tersedia. Pembersihan data digunakan memiliki kualitas yang baik dan tidak menimbulkan bias dalam proses pemodelan.

3.5.2 Transformasi dan Encoding Data

Dataset yang digunakan dalam penelitian ini mengandung data numerik dan data kategorikal. Oleh karena itu, dilakukan proses transformasi dan *encoding* agar seluruh data dapat diproses oleh algoritma *Random Forest*. Data kategorikal, seperti jenis tanaman, jenis tanah, dan kategori pH, diubah kedalam bentuk numerik menggunakan Teknik encoding. Selain itu, transformasi data juga dilakukan untuk menyeragamakan format data sehingga memudahkan proses analisis dan pelatihan model.

3.5.3 Normalisasi Data

Normalisasi data dilakukan untuk menyelaraskan rentang nilai antarvariabel, terutama pada parameter cuaca dan tanah yang memiliki skala berbeda. Meskipun *Random Forest* tidak terlalu sensitif terhadap perbedaan skala data, normalisasi tetap dipertimbangkan untuk menjaga konsistensi data dan memudahkan analisis lanjutan. Proses ini bertujuan agar setiap variabel memiliki kontribusi yang seimbang dalam proses pembelajaran model.

3.5.4 Transformasi Label Musim

Transformasi label musim merupakan salah satu tahap penting dalam proses pra-pemrosesan data yang bertujuan untuk menyesuaikan struktur dataset dengan kebutuhan penelitian. Dataset yang digunakan dalam penelitian ini memiliki variabel target *Season* yang terdiri dari empat kategori musim, yaitu *Spring*, *Summer*, *Fall*, dan *Winter*. Keempat kategori tersebut merepresentasikan pembagian musim yang umum digunakan pada wilayah beriklim subtropis.

Namun demikian, penelitian ini difokuskan pada pengembangan sistem rekomendasi musim tanam yang relevan dengan kondisi agroklimat wilayah tropis, khususnya Indonesia, yang pada umumnya hanya mengenal dua pola musim utama, yaitu musim kemarau (*Dry*) dan musim hujan (*Rainy*). Oleh karena itu, diperlukan proses transformasi label musim untuk menyederhanakan kategori musim dalam dataset agar sesuai dengan karakteristik iklim tropis yang menjadi fokus penelitian.

Proses transformasi dilakukan dengan memetakan kategori musim pada dataset ke dalam dua kelas utama. Kategori musim *Summer* diklasifikasikan sebagai musim kemarau (*Dry*) karena memiliki karakteristik suhu yang lebih tinggi serta kecenderungan curah hujan yang lebih rendah. Sementara itu, kategori musim *Spring*, *Fall*, dan *Winter* diklasifikasikan sebagai musim hujan (*Rainy*) karena memiliki karakteristik kondisi lingkungan yang relatif lebih lembap dengan intensitas curah hujan yang lebih tinggi dibandingkan musim kemarau.

Melalui proses transformasi label musim ini, dataset yang semula memiliki empat kategori musim kemudian disederhanakan menjadi dua kategori utama, yaitu *Dry* dan *Rainy*. Transformasi ini bertujuan untuk memastikan bahwa label target dalam dataset selaras dengan konteks agroklimat yang digunakan dalam penelitian,

sehingga model Random Forest dapat mempelajari pola hubungan antara parameter cuaca dan tanah dengan musim tanam secara lebih relevan dan representatif.

3.5.5 Pembagian Data

Setelah tahap pra-pemrosesan selesai, dataset dibagi menjadi dua bagian, yaitu data latih dan uji data. Data latih digunakan untuk melatih model *Random Forest*, sedangkan data uji digunakan untuk menguji kinerja model. Pembagian data dilakukan dengan perbandingan tertentu agar model dapat diuji secara objektif terhadap data yang belum pernah dilihat sebelumnya.

Dengan dilakukannya tahapan pra-pemrosesan data ini, dataset yang digunakan menjadi lebih bersih, terstruktur, dan siap digunakan dalam proses pelatihan serta pengujian model. Tahapan ini berperan penting dalam menghasilkan sistem rekomendasi musim tanam yang akurat dan andal.

3.6 Metode Random Forest

Metode *Random Forest* digunakan dalam penelitian ini untuk membangun sistem rekomendasi musim tanam tanaman hortikultura berdasarkan parameter cuaca dan tanah. *Random Forest* merupakan salah satu algoritma *machine learning* yang termasuk ke dalam metode *ensemble learning*, yaitu dengan menggabungkan banyak *decision tree* untuk menghasilkan prediksi yang lebih akurat dan stabil.

3.6.1 Konsep Dasar Random Forest

Random Forest bekerja dengan membangun sejumlah *decision tree* dari data latih yang diambil secara acak menggunakan teknik *bootstrap sampling*. Setiap *decision tree* menghasilkan prediksi, kemudian seluruh prediksi tersebut digabungkan untuk menentukan hasil akhir.

Pada kasus klasifikasi, hasil prediksi Random Forest ditentukan berdasarkan *majority voting*, yang secara matematis dapat dituliskan sebagai berikut :

$$\hat{y} = \text{mode}(f_1(x), f_2(x), \dots, f_B(x)) \quad (5)$$

Pendekatan ini membuat Random Forest mampu mengurangi variansi dan meningkatkan ketahanan model terhadap *overfitting*.

3.6.2 Kriteria Pemisahan Node (*Gini Impurity*)

Dalam proses pembentukan *decision tree*, *Random Forest* menggunakan ukuran *impurity* untuk menentukan pemisahan node terbaik. Salah satu ukuran yang umum digunakan adalah *Gini Impurity*, yang dirumuskan sebagai:

$$Gini = 1 - \sum_{i=1}^c p_i^2 \quad (6)$$

Nilai Gini yang lebih kecil menunjukkan pemisahan data yang lebih baik. Kriteria ini digunakan pada setiap *decision tree* untuk memilih atribut yang paling optimal dalam proses pembentukan model.

3.6.3 Parameter Random Forest

Model *Random Forest* dalam penelitian ini dikonfigurasi menggunakan beberapa parameter utama, yaitu:

1. Jumlah pohon (*n_estimators*), yang menunjukkan jumlah *decision tree* dalam mode.
2. Kedalaman maksimum pohon (*max_depth*), untuk membatasi kompleksitas model.
3. Jumlah fitur acak (*max_features*) yang digunakan pada setiap pemisahan node.

Pengaturan parameter ini bertujuan untuk memperoleh keseimbangan antara akurasi prediksi dan efisiensi komputasi.

3.6.4 Proses Pelatihan dan Prediksi Model

Proses pelatihan dilakukan dengan memasukkan data latih ke dalam model *Random Forest*. Model kemudian mempelajari hubungan antara variabel *input* berupa parameter cuaca dan tanah dengan variabel *output* berupa musim tanam tanaman hortikultura.

Setelah proses pelatihan selesai, model digunakan untuk melakukan prediksi terhadap data uji atau data baru. Prediksi yang dihasilkan selanjutnya digunakan sebagai rekomendasi musim tanam, sehingga dapat membantu pengguna dalam menentukan waktu tanam yang sesuai berdasarkan kondisi lingkungan yang diberikan.

3.6.5 Feature Importance

Random Forest juga mampu mengukur tingkat kepentingan setiap fitur dalam proses prediksi. Nilai *feature importance* dihitung berdasarkan rata-rata penurunan *impurity* yang disebabkan oleh suatu fitur pada seluruh *decision tree*, yang dirumuskan sebagai:

$$FI_j = \frac{1}{B} \sum_{b=1}^B \Delta I_b(j) \quad (7)$$

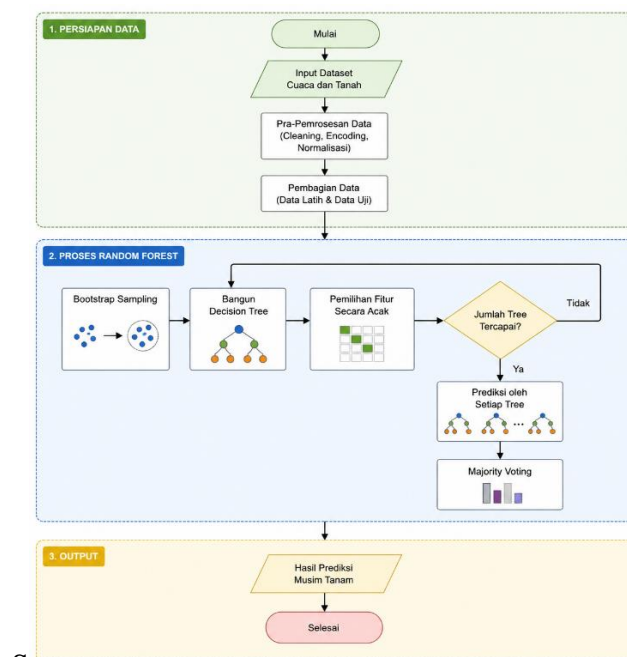
Informasi ini digunakan untuk mengetahui parameter cuaca dan tanah yang paling berpengaruh terhadap rekomendasi musim tanam.

3.6.6 Flowchart Random Forest

Flowchart proses *Random Forest* digunakan untuk menggambarkan tahapan kerja algoritma *Random Forest* dalam menghasilkan rekomendasi musim tanam tanaman hortikultura. Proses dimulai dengan memasukkan dataset parameter cuaca

dan tanah, kemudian dilakukan tahap pra-pemrosesan data untuk memastikan kualitas data. Dataset selanjutnya dibagi menjadi data latih dan data uji.

Pada tahap pelatihan, *Random Forest* membangun sejumlah *decision tree* menggunakan teknik *bootstrap sampling* dan pemilihan fitur secara acak. Proses ini dilakukan secara berulang hingga jumlah *decision tree* yang ditentukan tercapai. Setiap *decision tree* menghasilkan prediksi musim tanam, yang kemudian digabungkan menggunakan mekanisme *majority voting* untuk menentukan hasil prediksi akhir. Hasil prediksi tersebut digunakan sebagai rekomendasi musim tanam tanaman hortikultura.



Gambar 3.2 Flowchart Random Forest

3.7 Perancangan Sistem Rekomendasi

Perancangan sistem rekomendasi pada penelitian ini dilakukan untuk menggambarkan bagaimana komponen sistem saling berinteraksi dalam menghasilkan rekomendasi musim tanam tanaman hortikultura. Rancangan sistem mencakup aktor yang terlibat, alur input data, proses pengolahan data menggunakan

algoritma *Random Forest*, serta penyajian output rekomendasi kepada pengguna. Dengan adanya perancangan sistem yang jelas, proses implementasi dan evaluasi sistem dapat dilakukan secara terstruktur dan sistematis.

Perancangan sistem rekomendasi pada penelitian ini bertujuan untuk menghasilkan rekomendasi musim tanam tanaman hortikultura berdasarkan parameter cuaca dan tanah. Sistem dirancang agar mampu menerima data masukan, memproses data menggunakan model *Random Forest* yang telah dilatih, serta menghasilkan keluaran berupa rekomendasi musim tanam yang mudah dipahami oleh pengguna.

3.7.1 Arsitektur Sistem

Arsitektur sistem rekomendasi terdiri dari tiga komponen utama, yaitu input, proses, dan output. Pada bagian input, sistem menerima data parameter cuaca dan tanah yang meliputi suhu udara, curah hujan, kelembapan relatif, lama penyinaran, intensitas cahaya, pH tanah, jenis tanah, tingkat kesuburan, serta kandungan unsur hara nitrogen, fosfor, dan kalium. Data input tersebut disesuaikan dengan format dataset yang digunakan pada tahap pelatihan model.

Pada bagian proses, data input terlebih dahulu melalui tahap pra-pemrosesan yang sama dengan data latih, kemudian diproses oleh model *Random Forest* yang telah dilatih sebelumnya. Model ini melakukan prediksi berdasarkan pola hubungan yang telah dipelajari dari data historis.

Pada bagian output, sistem menampilkan hasil prediksi dalam bentuk rekomendasi musim tanam tanaman hortikultura. Rekomendasi ini diharapkan dapat membantu pengguna dalam menentukan waktu tanam yang sesuai berdasarkan kondisi lingkungan yang diberikan.

3.7.2 Alur Kerja Sistem Rekomendasi

Alur kerja sistem rekomendasi dimulai ketika pengguna memasukkan parameter cuaca dan tanah ke dalam sistem. Data tersebut kemudian divalidasi dan diproses agar sesuai dengan format data latih. Selanjutnya, data dimasukkan ke dalam model Random Forest yang telah dilatih untuk menghasilkan prediksi musim tanam. Hasil prediksi ini kemudian disajikan kepada pengguna dalam bentuk rekomendasi musim tanam yang sesuai dengan kondisi lingkungan yang diberikan.

Model *Random Forest* menghasilkan prediksi musim tanam berdasarkan hasil *majority voting* dari seluruh *decision tree*. Hasil prediksi ini kemudian ditampilkan sebagai rekomendasi musim tanam tanaman hortikultura. Dengan alur kerja ini, sistem rekomendasi dapat memberikan saran musim tanam secara cepat dan konsisten.

3.7.3 Keluaran Sistem (Output Rekomendasi)

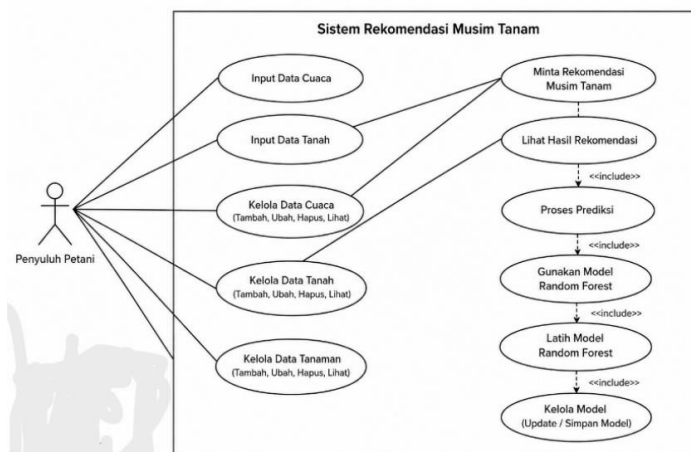
Keluaran dari sistem rekomendasi berupa informasi musim tanam yang direkomendasikan untuk tanaman hortikultura tertentu berdasarkan kondisi cuaca dan tanah. Rekomendasi ini disajikan dalam bentuk yang sederhana dan informatif agar mudah dipahami oleh pengguna. Selain rekomendasi musim tanam, sistem juga dapat menampilkan informasi pendukung, seperti tingkat keyakinan prediksi atau parameter lingkungan yang paling berpengaruh terhadap hasil rekomendasi.

3.7.4 Use Case Diagram Sistem Rekomendasi Musim Tanam

Use Case Diagram digunakan untuk menggambarkan fungsionalitas sistem rekomendasi musim tanam tanaman hortikultura serta interaksi antara aktor dan sistem. Diagram ini menunjukkan layanan-layanan utama yang disediakan oleh sistem, mulai dari proses manajemen data lingkungan hingga penyajian hasil

rekomendasi musim tanam kepada pengguna. Dengan Use Case Diagram, ruang lingkup sistem dapat dipahami secara jelas dari sudut pandang pengguna yang kini difokuskan pada peran tenaga ahli di lapangan.

Pada Use Case Diagram ini terdapat satu aktor utama, yaitu Penyuluh Petani. Aktor Penyuluh Petani memiliki kewenangan penuh dalam mengoperasikan sistem, yang meliputi penginputan data cuaca dan tanah, pengelolaan data tanaman hortikultura, serta melakukan pelatihan dan pengelolaan model *Random Forest*. Selain itu, Penyuluh Petani juga bertugas mengajukan permintaan rekomendasi dan melihat hasil prediksi yang dihasilkan oleh sistem untuk kemudian disosialisasikan kepada petani. Sistem akan menggunakan model *Random Forest* untuk melakukan proses prediksi musim tanam berdasarkan data-data parameter lingkungan yang telah diinputkan oleh penyuluh. Dengan adanya pembaruan pada Use Case Diagram ini, alur fungsional sistem menjadi lebih terpusat dan memudahkan proses pemantauan serta pengambilan keputusan berbasis data agroklimat.



Gambar 3.3 Use Case Diagram Sistem Rekomendasi Musim Tanam

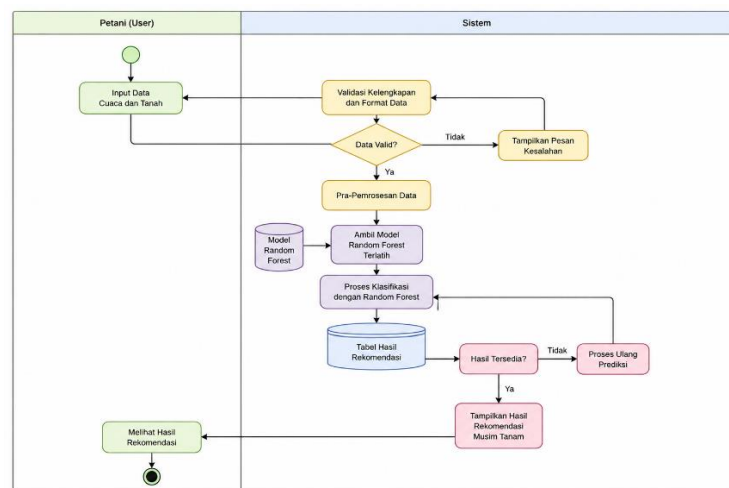
3.7.5 Activity Diagram Sistem Rekomendasi Musim Tanam

Activity Diagram digunakan untuk menggambarkan alur aktivitas sistem rekomendasi musim tanam tanaman hortikultura berdasarkan parameter cuaca dan

tanah. Diagram ini menunjukkan interaksi sistematis antara aktor Penyuluh Petani dan Sistem dalam rangkaian proses mulai dari input data, validasi otomatis, pra-pemrosesan, klasifikasi menggunakan algoritma *Random Forest*, hingga penyajian hasil rekomendasi.

Proses pada *Activity Diagram* dimulai saat Penyuluh Petani melakukan input data parameter cuaca dan tanah ke dalam sistem. Setelah data dimasukkan, sistem secara otomatis melakukan validasi kelengkapan dan format data untuk menjamin kualitas informasi yang akan diolah. Apabila data ditemukan tidak valid, sistem akan menampilkan pesan kesalahan dan mengarahkan Penyuluh Petani untuk melakukan input ulang pada form yang tersedia.

Jika data dinyatakan valid, sistem akan melanjutkan ke tahap pra-pemrosesan data dan memanggil model *Random Forest* yang telah dilatih sebelumnya dan disimpan dalam format serialisasi .pkl. Data tersebut kemudian diproses melalui mekanisme klasifikasi untuk menghasilkan prediksi musim tanam yang paling sesuai bagi tanaman hortikultura. Hasil prediksi tersebut disimpan ke dalam tabel hasil rekomendasi dan ditampilkan melalui antarmuka sistem. Proses berakhir setelah hasil rekomendasi berhasil ditampilkan dan dilihat oleh Penyuluh Petani.

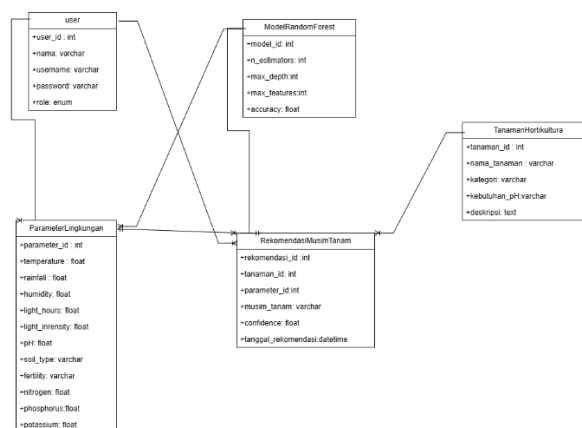


Gambar 3.4 Activity Diagram Sistem Rekomendasi Musim Tanam

3.7.6 Class Diagram Sistem Rekomendasi Musim Tanam

Class Diagram digunakan untuk menggambarkan struktur kelas dan hubungan antar kelas yang membentuk Sistem Rekomendasi Musim Tanam Tanaman Hortikultura berdasarkan Parameter Cuaca dan Tanah menggunakan *Ensemble Learning*. Diagram ini menunjukkan entitas utama yang terlibat dalam sistem, yaitu *User*, Tanaman Hortikultura, Parameter Lingkungan, Model *Random Forest*, dan Rekomendasi Musim Tanam. Setiap kelas merepresentasikan objek nyata dalam sistem yang memiliki atribut dan relasi tertentu sesuai dengan fungsinya masing-masing.

Hubungan antar kelas pada *Class Diagram* menunjukkan bahwa *User* berperan dalam memasukkan data parameter cuaca dan tanah ke dalam sistem. Data tersebut disimpan dalam kelas Parameter Lingkungan dan selanjutnya diproses oleh kelas Model *Random Forest* yang berfungsi sebagai mesin klasifikasi berbasis *ensemble learning*. Proses klasifikasi menghasilkan keluaran berupa rekomendasi musim tanam yang disimpan dalam kelas Rekomendasi Musim Tanam dan dikaitkan dengan kelas Tanaman Hortikultura. Dengan demikian, *Class Diagram* ini menjadi dasar perancangan basis data serta mendukung implementasi sistem rekomendasi musim tanam secara terstruktur dan konsisten.

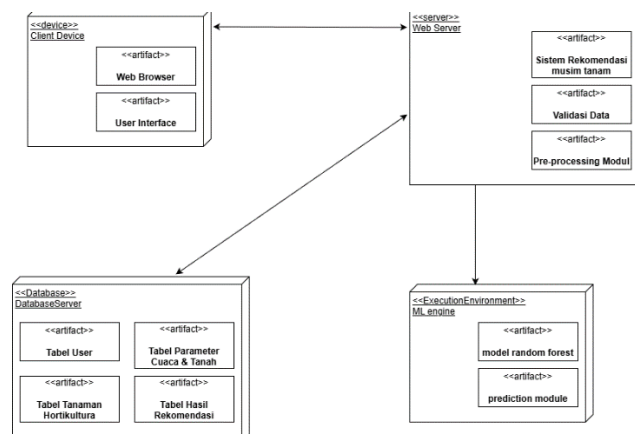


Gambar 3.5 *Class Diagram* Sistem Rekomendasi Musim Tanam

3.7.7 Deployment Diagram Sistem Rekomendasi Musim Tanam

Deployment Diagram digunakan untuk menggambarkan arsitektur fisik dan penempatan komponen perangkat lunak pada Sistem Rekomendasi Musim Tanam Tanaman Hortikultura. Diagram ini menunjukkan bagaimana sistem diimplementasikan pada lingkungan nyata, mulai dari perangkat pengguna hingga server aplikasi, komponen pemrosesan *machine learning*, dan basis data yang digunakan untuk menyimpan informasi sistem.

Pada *Deployment Diagram* ini, pengguna mengakses sistem melalui perangkat client menggunakan *web browser*. Permintaan dari *client* diteruskan ke *web server* yang berfungsi sebagai *application server* untuk menangani proses validasi data dan pengelolaan alur sistem. *Web server* selanjutnya berkomunikasi dengan komponen *Machine Learning Engine* yang menjalankan model *Random Forest* terlatih untuk melakukan prediksi musim tanam berdasarkan parameter cuaca dan tanah. Hasil prediksi kemudian disimpan ke dalam *database server* dan dikirim kembali ke *web server* untuk ditampilkan kepada pengguna. Dengan demikian, *Deployment Diagram* ini memberikan gambaran yang jelas mengenai hubungan dan komunikasi antar komponen sistem dalam mendukung proses rekomendasi musim tanam.



Gambar 3.6 *Deployment Diagram* Sistem Rekomendasi Musim Tanam

3.7.8 Perancangan Wireframe Antarmuka Sistem

Wireframe antarmuka sistem digunakan untuk menggambarkan rancangan awal tampilan website "Asisten Digital Penyuluh Pertanian" sebelum tahap implementasi dilakukan, dengan tujuan memastikan fungsi sistem dan kebutuhan pengguna terakomodasi dengan baik. Rancangan ini disusun dalam bentuk *low fidelity* yang berfokus pada struktur halaman, tata letak komponen, serta alur interaksi pengguna tanpa menitikberatkan pada aspek estetika desain. Secara hierarkis, sistem dimulai dari halaman Dashboard yang berfungsi sebagai pusat navigasi utama dan memberikan gambaran singkat mengenai fungsi sistem kepada penyuluh pertanian.

Proses inti sistem dilakukan pada halaman Analisis & Rekomendasi, di mana pengguna memasukkan berbagai parameter cuaca seperti suhu, curah hujan, kelembapan, lama penyinaran, dan intensitas cahaya, serta parameter tanah yang mencakup pH tanah, jenis tanah, kesuburan, dan kandungan unsur hara (Nitrogen, Fosfor, Kalium). Data lingkungan tersebut menjadi dasar utama dalam proses klasifikasi musim tanam menggunakan algoritma Ensemble Learning (Random Forest). Selanjutnya, halaman Hasil Analisis SPK akan menampilkan keluaran berupa prediksi musim tanam, tingkat akurasi hasil prediksi, serta daftar rekomendasi komoditas tanaman hortikultura yang paling sesuai dengan kondisi lahan yang telah diinput.

Untuk mendukung manajemen data dan transparansi, sistem dilengkapi dengan halaman Riwayat Analisis yang menyajikan log lengkap dari setiap proses klasifikasi yang telah dilakukan sebelumnya. Selain itu, tersedia halaman Manajemen Data yang memfasilitasi pengunggahan dataset baru dalam format CSV

Asisten Digital
Penyuluh Pertanian

- Dashboard
- Analisis & Rekomendasi
- Riwayat Analisis**
- Manajemen Data
- Data Hortikultura
- Pusat Pelatihan Model

© 2025 SPK Penyuluh

Riwayat Analisis SPK

Daftar lengkap semua analisis dan rekomendasi yang telah dilakukan oleh sistem.

WAKTU	MUSIM PREDIKSI	REKOMENDASI	CONF.	SUHU	HUJAN
08-03-2026 17:48	Summer	Broccoli, Cabbage, Chard, Chili Peppers, Cucumbers, Eggplants, Gra...	44.2%	19.0°C	600.0 mm
07-03-2026 19:25	Summer	Broccoli, Cabbage, Chard, Chili Peppers, Cucumbers, Eggplants, Gra...	50.9%	23.0°C	700.0 mm
07-03-2026 17:49	Summer	Broccoli, Cabbage, Chard, Chili Peppers, Cucumbers, Eggplants, Gra...	42.5%	20.0°C	400.0 mm
03-03-2026 17:39	Summer	Broccoli, Cabbage, Chard, Chili Peppers, Cucumbers, Eggplants, Gra...	43.5%	30.0°C	40.0 mm
03-03-2026 19:38	Summer	Broccoli, Cabbage, Chard, Chili Peppers, Cucumbers, Eggplants, Gra...	37.7%	33.0°C	300.0 mm
02-03-2026 17:44	Fall	Arugula, Beet, Broccoli, Cabbage, Cauliflowers, Chard, Chili Peppers...	47.9%	23.4°C	400.0 mm
02-03-2026 17:47	Summer	Broccoli, Cabbage, Chard, Chili Peppers, Cucumbers, Eggplants, Gra...	47.9%	24.8°C	400.0 mm
02-03-2026 17:37	Summer	Broccoli, Cabbage, Chard, Chili Peppers, Cucumbers, Eggplants, Gra...	41.0%	24.8°C	300.0 mm

Gambar 3.10 Wireframe Halaman Riwayat Analisis

Asisten Digital
Penyuluh Pertanian

- Dashboard
- Analisis & Rekomendasi
- Riwayat Analisis
- Manajemen Data**
- Data Hortikultura
- Pusat Pelatihan Model

© 2025 SPK Penyuluh

Upload & Latih Ulang Model

Unggah file Excel atau CSV untuk memperbarui dataset dan melatih ulang model secara otomatis.

Pilih File Dataset (.csv, .xlsx, .xls)

Klik atau seret file ke sini!

Andakan file memiliki kolom: temperature, pestik, humidity, soil_type, season, etc

[Upload & Latih](#)

Riwayat Upload Dataset			
NAMA FILE	TANGGAL	JUMLAH DATA	STATUS
dataset_bersih_1000_decimal.csv.xlsx	02-03-2026 17:48	1000 Baris	Berhasil Diunggah
dataset_bersih_1000_int4.csv	02-03-2026 17:43	1000 Baris	Berhasil Diunggah

RINCIAN PARAMETER INPUT v

[← Kembali ke Dashboard](#)
[Mulai Analisis Baru](#)

Gambar 3.11 Wireframe Halaman Management Data

Asisten Digital
Penyuluh Pertanian

- Dashboard
- Analisis & Rekomendasi
- Riwayat Analisis
- Manajemen Data
- Data Hortikultura
- Pusat Pelatihan Model**

© 2025 SPK Penyuluh

Pusat Pelatihan & Evaluasi

Pantau kinerja model cerdas SPK dan latih ulang dengan data terbaru. [Latih Ulang Sekarang](#)

AKURASI

61.3%

Kategori prediksi keseluruhan

PREKSI

59.9%

Tingkat positif yang benar

RECALL

61.3%

Sensitivitas model

F1-SCORE

60.3%

Keseimbangan presisi & recall

Riwayat Pelatihan			
TANGGAL	JUMLAH DATA	AKURASI	F1-SCORE
02-03-2026 17:48	2000	61.29%	60.28%
02-03-2026 17:43	1000	50.00%	46.73%

Gambar 3.12 Wireframe Halaman Data Holtikultura

Asisten Digital Penyuluh Pertanian		Dataset Pelatihan										
Dashboard Analisis & Rekomendasi Riwayat Analisis Manajemen Data Data Hortikultura Pusat Pelatihan Model		ID	SEASON (TARGET)	SUHU	KEJARAN	LEMBAB	CAHAYA (JAM)	INTENSITAS	PH	TANAH	KESUBURAN	N-P-K
		#2000	Summer	24.8	1112.7	87.4	9.0	435	6.1	Sandy Loam	High	48-21-75
		#1999	Fall	18.5	1183.7	87.6	12.2	388	6.6	Sandy Loam	Medium	138-84-224
		#1998	Fall	22.7	702.4	86.6	6.7	361	5.8	Loam	Medium	174-94-184
		#1987	Spring	19.7	888.8	76.8	6.4	359	5.8	Loam	High	151-73-188
		#1986	Summer	20.0	781.7	91.6	13.5	630	6.9	Loam	Medium	144-67-188
		#1995	Spring	24.6	1396.5	55.4	8.7	800	6.9	Sandy Loam	High	150-154-133
		#1984	Fall	21.7	712.7	45.0	12.5	931	6.3	Loam	High	170-252-284
		#1983	Spring	17.8	830.8	52.8	12.1	220	6.4	Sandy Loam	High	121-88-59
		#1982	Summer	22.6	470.9	50.9	7.2	557	5.7	Sandy Loam	High	188-180-220
#1981	Fall	18.5	1298.8	54.9	13.0	456	6.3	Sandy Loam	Medium	141-49-210		
#1990	Summer	35.3	558.1	80.5	7.0	313	6.8	Sandy Loam	High	84-35-51		
#1889	Spring	22.7	614.0	89.8	6.9	313	6.3	Sandy Loam	High	95-44-44		
#1888	Spring	28.2	886.8	88.7	6.9	327	6.2	Sandy Loam	High	96-48-28		
#1887	Spring	33.9	482.5	89.9	7.4	338	6.3	Sandy Loam	High	86-45-21		
#1886	Summer	24.6	482.5	90.3	7.4	338	6.4	Sandy Loam	High	88-32-51		
#1885	Spring	24.6	682.7	90.3	6.9	301	6.1	Sandy Loam	High	84-34-50		
#1884	Summer	33.2	671.2	90.0	7.2	309	6.5	Sandy Loam	High	91-53-52		
#1883	Spring	31.8	619.6	90.0	7.7	302	6.5	Sandy Loam	High	86-37-52		

Gambar 3.13 Wireframe Halaman Pusat Pelatihan Model

3.8 Metode Evaluasi Model

Metode evaluasi digunakan untuk mengukur kinerja model *Random Forest* dalam menghasilkan rekomendasi musim tanam tanaman hortikultura. Evaluasi dilakukan dengan membandingkan hasil prediksi model terhadap data aktual pada data uji. Metrik evaluasi yang digunakan dalam penelitian ini meliputi *accuracy*, *precision*, *recall*, dan *F1-score*, yang umum digunakan pada permasalahan klasifikasi.

3.8.1 Confusion Matrix

Confusion matrix digunakan untuk menggambarkan perbandingan antara hasil prediksi model dan data aktual. *Confusion matrix* terdiri dari empat komponen utama, yaitu:

- True Positive* (TP) : data positif yang diprediksi benar
- True Negative* (TN) : data negatif yang diprediksi benar
- False Positive* (FP) : data negatif yang diprediksi sebagai positif
- False Negative* (FN) : data positif yang diprediski sebagai negatif

Confusion matrix menjadi dasar perhitungan metrik evaluasi lainnya.

3.8.2 Accuracy

Accuracy digunakan untuk mengukur tingkat ketepatan model dalam melakukan klasifikasi secara keseluruhan.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (8)$$

3.8.3 Precision

Precision digunakan untuk mengukur tingkat ketepatan prediksi positif yang dihasilkan oleh model.

$$Precision = \frac{TP}{TP + FP} \quad (9)$$

3.8.4 Recall

Recall digunakan untuk mengukur kemampuan model dalam menemukan seluruh data positif yang sebenarnya.

$$Recall = \frac{TP}{TP + FN} \quad (10)$$

3.8.5 F1-Score

F1-Score merupakan rata-rata harmonik dari *precision* dan *recall* yang digunakan untuk menyeimbangkan kedua metrik tersebut.

$$F1 - Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (11)$$

Penggunaan beberapa metrik evaluasi bertujuan untuk memberikan gambaran yang lebih komprehensif mengenai kinerja model. *Accuracy* menunjukkan performa model secara umum, sedangkan *precision* dan *recall* memberikan informasi lebih detail mengenai kemampuan model dalam memprediksi kelas

tertentu. *F1-score* digunakan sebagai metrik gabungan untuk menilai keseimbangan antara *precision* dan *recall*.

3.9 Jadwal Penelitian

Penelitian ini dilaksanakan dalam jangka waktu 3 bulan, dimulai dari tahap persiapan hingga penyusunan laporan akhir. Penyusunan jadwal penelitian bertujuan untuk memastikan setiap tahapan penelitian dapat dilaksanakan secara sistematis dan terstruktur, sesuai dengan alur metodologi yang telah dijelaskan pada subbab sebelumnya. Rincian jadwal penelitian disajikan pada Tabel 3.1.

Tabel 3.2 Jadwal Penelitian

No	Tahapan Penelitian	Bulan 1	Bulan 2	Bulan 3
1	Studi Literatur dan Penyusunan Proposal			
2	Pengumpulan Dataset			
3	Pra-Pemrosesan Data			
4	Pembangunan Model Random Forest			
5	Pengujian dan Evaluasi			
6	Analisis Hasil Penelitian			
7	Penyusunan Laporan Skripsi			

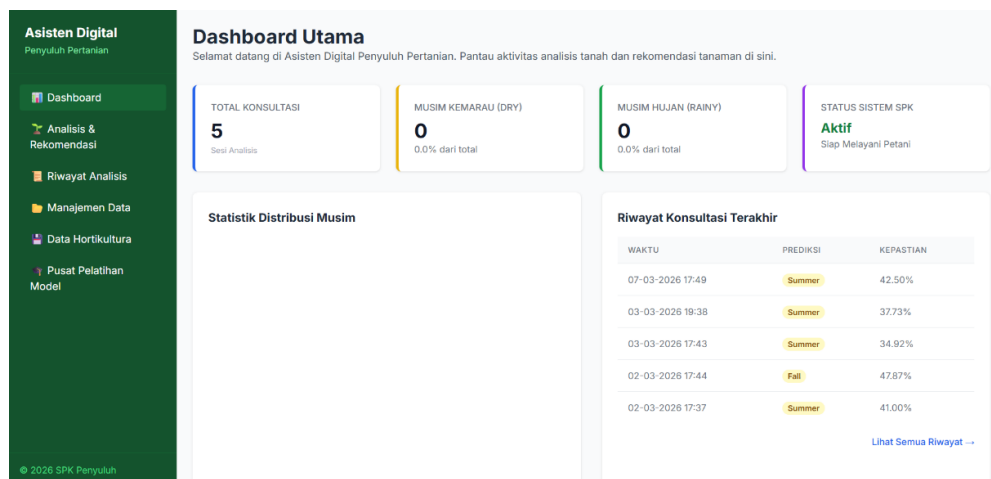
BAB IV

HASIL DAN PEMBAHASAN

4.1 Implementasi Antarmuka

4.1.1 Implementasi Halaman Dashboard

Implementasi antarmuka pada halaman dashboard merupakan komponen vital dalam sistem ini karena berfungsi sebagai pusat kendali informasi dan pemantauan bagi penyuluh pertanian. Sebagai halaman utama setelah proses autentikasi, dashboard dirancang untuk memberikan gambaran komprehensif mengenai aktivitas klasifikasi yang telah dilakukan oleh sistem secara keseluruhan. Secara teknis, halaman ini diimplementasikan melalui rute utama pada berkas `app.py` yang melakukan kueri data dari tabel `PredictionLog` dalam basis data `SQLite` untuk menghasilkan statistik deskriptif secara otomatis. Informasi yang dikumpulkan mencakup total sesi konsultasi, jumlah prediksi musim kemarau, dan jumlah prediksi musim hujan, yang kemudian dikirimkan ke template `dashboard.html` untuk dirender menjadi elemen visual yang informatif.



Gambar 4.1 Halaman Dashboard

Struktur visual dashboard dibagi menjadi beberapa bagian utama, dimulai dengan deretan kartu statistik (*stats cards*) yang menyajikan indikator kinerja utama sistem. Kartu pertama menampilkan total konsultasi yang merepresentasikan seluruh sesi analisis yang pernah dijalankan oleh penyuluh, memberikan gambaran mengenai intensitas penggunaan sistem. Kartu-kartu berikutnya menampilkan distribusi frekuensi klasifikasi untuk musim kemarau (*Dry*) dan musim hujan (*Rainy*), lengkap dengan persentase distribusinya terhadap total data.

Secara akademis, penyajian persentase ini penting untuk memberikan wawasan makro mengenai tren iklim yang terekam dalam sistem. Selain itu, terdapat kartu status sistem yang menunjukkan indikator "Aktif", memberikan kepastian teknis bahwa mesin Sistem Pendukung Keputusan (SPK) siap melayani kebutuhan petani secara *real-time*.

Guna memperkuat representasi data, sistem mengintegrasikan visualisasi grafis berupa diagram donat (*doughnut chart*) yang terletak pada sisi kiri halaman. Implementasi grafik ini menggunakan pustaka JavaScript Chart.js yang diatur untuk melakukan pemetaan data statistik dari backend Flask ke dalam elemen kanvas HTML. Penggunaan warna kuning untuk merepresentasikan musim kemarau dan warna hijau untuk musim hujan dilakukan secara sengaja untuk memberikan isyarat visual yang selaras dengan kondisi alamiah, sehingga memudahkan penyuluh dalam menginterpretasikan distribusi musim secara cepat tanpa harus membaca angka secara mendetail.

Visualisasi ini tidak hanya bersifat estetis, tetapi juga berfungsi sebagai instrumen analisis tren temporal bagi pengguna sistem. Di samping visualisasi grafis, dashboard juga menyajikan tabel riwayat konsultasi terakhir pada sisi kanan

untuk memberikan transparansi terhadap aktivitas terbaru. Tabel ini menampilkan lima rekaman data terakhir yang mencakup stempel waktu (timestamp), label hasil prediksi musim, dan tingkat kepastian (confidence level) dari model Random Forest. Penampilan tingkat kepastian dalam bentuk persentase merupakan aspek krusial dalam sistem berbasis machine learning karena memberikan informasi mengenai derajat keyakinan model terhadap input parameter cuaca dan tanah yang diberikan. Melalui tabel ini, penyuluh dapat segera meninjau hasil analisis terakhir secara kronologis, lengkap dengan tautan untuk mengakses seluruh riwayat konsultasi jika diperlukan analisis yang lebih mendalam pada halaman histori.

Secara keseluruhan, implementasi halaman dashboard ini mencerminkan integrasi yang erat antara logika *backend* dan presentasi *frontend*. Penggunaan framework CSS Tailwind dalam `dashboard.html` memastikan tata letak yang bersih dan profesional, sementara logika asinkron pada Flask memastikan bahwa data yang ditampilkan selalu akurat sesuai dengan kondisi terkini dalam basis data. Dashboard ini tidak hanya sekadar antarmuka statis, melainkan sebuah platform pemantauan dinamis yang mendukung efektivitas kerja penyuluh pertanian dalam memantau parameter lingkungan dan memberikan rekomendasi masa tanam yang tepat berdasarkan data empiris yang terhimpun dalam sistem.

4.1.2 Implementasi Halaman Musim Tanam

Halaman Klasifikasi Musim Tanam merupakan antarmuka interaktif utama di mana pengguna, khususnya penyuluh pertanian, memasukkan parameter fisik lingkungan untuk dianalisis oleh model *machine learning*. Secara teknis, halaman ini diimplementasikan melalui rute `/predict` dengan metode permintaan GET pada framework Flask.

Analisis Lahan & Rekomendasi
Masukkan parameter cuaca dan tanah lapangan untuk mendapatkan rekomendasi tanaman yang presisi.

Parameter Cuaca

Suhu (°C) Curah Hujan (mm)

Kelembaban (%) Lama Penyinaran (jam)

Intensitas Cahaya (lux)

Parameter Tanah

pH Tanah Jenis Tanah Kesuburan

Nitrogen (N) Fosfor (P) Kalium (K)

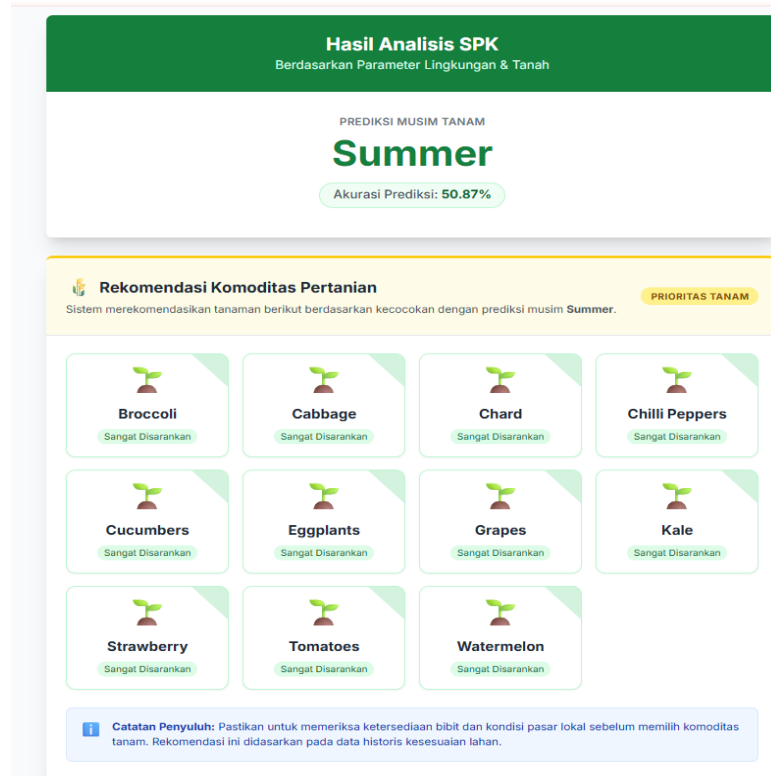
Gambar 4.2 Halaman Analisis Lahan & Rekomendasi

Antarmuka ini dirancang menggunakan formulir input yang komprehensif, mencakup variabel klimatologi seperti suhu, curah hujan, kelembapan, durasi penyinaran matahari, dan intensitas cahaya, serta variabel edafik atau kondisi tanah yang meliputi tingkat keasaman (pH), kadar Nitrogen (N), Fosfor (P), dan Kalium (K). Selain input numerik, halaman ini juga menyediakan pilihan kategorikal untuk jenis tanah (*soil type*) dan tingkat kesuburan (*fertility*) guna memastikan data yang dimasukkan selaras dengan skema data yang dipahami oleh model *Random Forest*.

Implementasi formulir pada halaman ini mengedepankan aspek fungsionalitas dan akurasi data. Setiap elemen input dilengkapi dengan validasi tipe data guna meminimalisir kesalahan pengguna (*human error*) sebelum data dikirim ke server untuk diproses. Secara akademis, penyediaan parameter yang mendetail ini menunjukkan bahwa sistem tidak hanya melakukan klasifikasi berdasarkan variabel tunggal, melainkan mempertimbangkan interaksi multidimensi antara kondisi atmosfer dan nutrisi tanah. Hal ini krusial dalam konteks hortikultura, di

mana kesesuaian musim tanam sangat bergantung pada ketersediaan unsur hara dan kecukupan air yang direpresentasikan melalui variabel input tersebut.

4.1.3 Implementasi Hasil Prediksi dan Rekomendasi



Gambar 4.3 Hasil Prediksi dan Rekomendasi

Halaman hasil prediksi merupakan komponen krusial dalam antarmuka sistem yang berfungsi untuk menyajikan output dari pemrosesan model *machine learning* secara transparan dan informatif kepada pengguna. Implementasi halaman ini dilakukan melalui rute `/predict` dengan metode POST pada berkas `app.py`, di mana setelah data parameter cuaca dan tanah diterima, sistem melakukan serangkaian operasi komputasi untuk menghasilkan klasifikasi akhir. Secara teknis, model *Random Forest* yang telah dimuat (`rf_model`) memproses data input yang telah dinormalisasi untuk menghasilkan indeks prediksi, yang kemudian dikonversi kembali menjadi label kategori musim (seperti "Rainy" atau "Dry") menggunakan

objek `le_season`. Hasil ini kemudian dirender ke dalam `template result.html` bersama dengan data input asli dan nilai kepastian prediksi.

Aspek penting yang ditampilkan pada halaman ini adalah nilai kepastian atau *confidence level* dari hasil klasifikasi. Nilai ini diperoleh dengan mengekstraksi probabilitas prediksi menggunakan fungsi `predict_proba` dari model *Random Forest*, di mana sistem mengambil nilai probabilitas tertinggi dari seluruh kelas yang ada dan mengonversinya ke dalam satuan persentase. Penyajian nilai kepastian ini dalam paragraf akademik dipandang sebagai bentuk transparansi model, yang memungkinkan penyuluh pertanian untuk mengukur tingkat risiko dan validitas dari rekomendasi yang diberikan oleh sistem. Secara visual, hasil prediksi ini ditampilkan dengan penekanan pada label musim menggunakan elemen grafis yang responsif, sehingga pengguna dapat segera mengidentifikasi musim tanam yang terdeteksi berdasarkan data lingkungan yang dimasukkan.

Selain hasil klasifikasi musim, halaman ini juga mengintegrasikan sistem rekomendasi tanaman hortikultura yang bersifat dinamis. Proses ini melibatkan pemetaan antara label musim hasil prediksi dengan basis data rekomendasi yang tersimpan dalam variabel `recommendation_df`. Logika yang diimplementasikan dalam berkas `app.py` melakukan filter terhadap dataset tersebut berdasarkan musim yang sesuai; misalnya, jika model memprediksi musim "Dry", sistem akan secara otomatis mencari daftar tanaman yang secara agronomis cocok ditanam pada kondisi tersebut. Daftar tanaman hortikultura ini kemudian ditampilkan dalam bentuk daftar yang sistematis pada bagian bawah halaman hasil, memberikan panduan praktis bagi petani mengenai jenis komoditas yang paling optimal untuk ditanam.

Secara sistematis, seluruh hasil yang ditampilkan pada halaman ini juga dicatat secara otomatis ke dalam tabel `PredictionLog` pada basis data SQLite. Data yang disimpan mencakup seluruh parameter input (seperti suhu, curah hujan, dan pH tanah), label musim hasil prediksi, nilai kepastian, hingga daftar rekomendasi tanaman yang diberikan. Implementasi pencatatan otomatis ini memastikan bahwa setiap sesi analisis yang dilakukan melalui halaman hasil prediksi memiliki rekam jejak digital yang permanen, sehingga dapat ditinjau kembali di masa mendatang untuk kebutuhan audit atau pemantauan akurasi hasil tanam secara riil di lapangan.

4.1.4 Implementasi Riwayat Analisis

Riwayat Analisis SPK
Daftar lengkap semua analisis dan rekomendasi yang telah dilakukan oleh sistem.

WAKTU	MUSIM PREDIKSI	REKOMENDASI	CONF.	SUHU	HUJAN
07-03-2026 19:25	Summer	Broccoli, Cabbage, Chard, Chilli Peppers, Cucumbers, Eggpla...	50.9%	23.0°C	700.0 mm
07-03-2026 17:49	Summer	Broccoli, Cabbage, Chard, Chilli Peppers, Cucumbers, Eggpla...	42.5%	20.0°C	400.0 mm
03-03-2026 19:38	Summer	Broccoli, Cabbage, Chard, Chilli Peppers, Cucumbers, Eggpla...	37.7%	33.0°C	40.0 mm
03-03-2026 17:43	Summer	Broccoli, Cabbage, Chard, Chilli Peppers, Cucumbers, Eggpla...	34.9%	24.8°C	300.0 mm
02-03-2026 17:44	Fall	Arugula, Beet, Broccoli, Cabbage, Cauliflowers, Chard, Chilli P...	47.9%	23.4°C	400.0 mm
02-03-2026 17:37	Summer	Broccoli, Cabbage, Chard, Chilli Peppers, Cucumbers, Eggpla...	41.0%	24.8°C	300.0 mm

Gambar 4.4 Halaman Riwayat Analisis

Implementasi halaman riwayat analisis merupakan perwujudan dari fungsi transparansi dan audit sistem, yang memungkinkan pengguna untuk meninjau kembali seluruh rekam jejak klasifikasi yang telah dilakukan sebelumnya. Secara teknis, fitur ini dikembangkan melalui rute `/history` pada framework Flask yang mengandalkan logika penarikan data dari basis data SQLite menggunakan ORM SQLAlchemy. Fungsi ini melakukan kueri secara menyeluruh pada tabel `PredictionLog` dengan menerapkan pengurutan berdasarkan stempel waktu secara menurun (*descending*), sehingga informasi yang paling mutakhir selalu

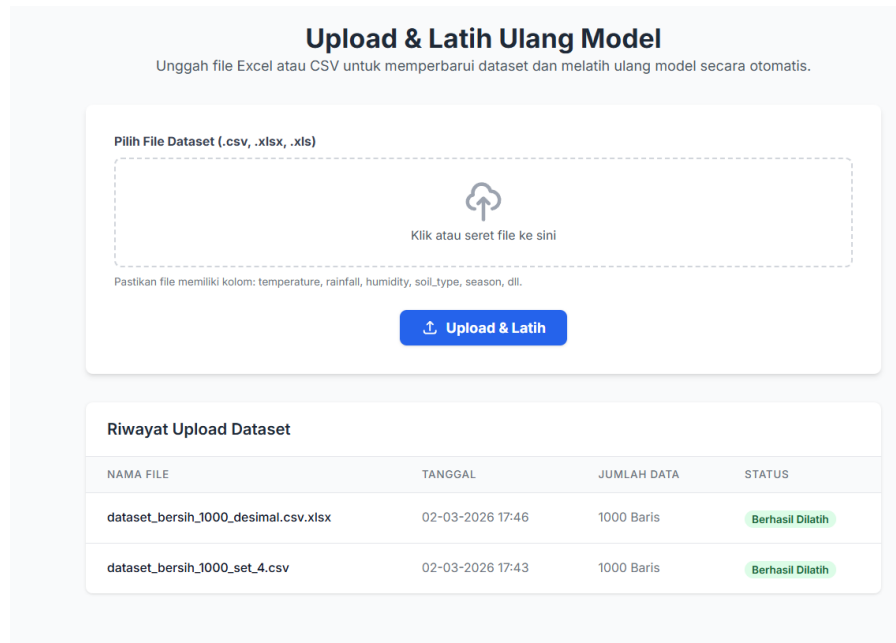
disajikan pada baris teratas tabel. Hal ini sangat krusial dalam konteks manajemen pertanian hortikultura, karena memungkinkan penyuluh untuk memantau perubahan rekomendasi seiring dengan fluktuasi parameter lingkungan yang dimasukkan dari waktu ke waktu.

Penyajian data pada antarmuka riwayat diimplementasikan menggunakan struktur tabel yang responsif dan sistematis melalui berkas `history.html`. Tabel tersebut mencakup informasi mendalam yang meliputi waktu pelaksanaan analisis, hasil prediksi musim, daftar rekomendasi tanaman, tingkat kepastian (*confidence level*), serta parameter cuaca utama seperti suhu dan curah hujan. Secara visual, hasil klasifikasi diberikan penanda warna yang kontras, di mana kategori musim hujan (*Rainy*) ditandai dengan label hijau dan musim kemarau (*Dry*) dengan label kuning guna mempercepat proses identifikasi kualitatif oleh pengguna. Selain itu, sistem menyertakan fitur *tooltip* pada kolom rekomendasi untuk menangani teks yang panjang, memastikan bahwa antarmuka tetap rapi tanpa menghilangkan detail informasi komoditas yang disarankan.

Keberadaan modul riwayat ini memberikan nilai tambah berupa kemampuan pelacakan (*traceability*) terhadap setiap keputusan yang diambil oleh Sistem Pendukung Keputusan (SPK). Setiap baris data dalam riwayat mencerminkan kondisi agroklimat spesifik yang pernah diproses oleh model *Random Forest*, sehingga dapat digunakan sebagai bahan evaluasi jika terjadi ketidaksesuaian antara hasil prediksi dengan kondisi riil di lapangan. Integrasi antara penyimpanan permanen pada database dengan penyajian data yang terorganisir di sisi *frontend* memastikan bahwa sistem tidak hanya berfungsi sebagai alat prediksi sesaat, melainkan juga sebagai basis pengetahuan (*knowledge base*) yang menyimpan

pola-pola klasifikasi musim tanam hortikultura yang telah diuji secara empiris oleh pengguna.

4.1.5 Implementasi Management Data



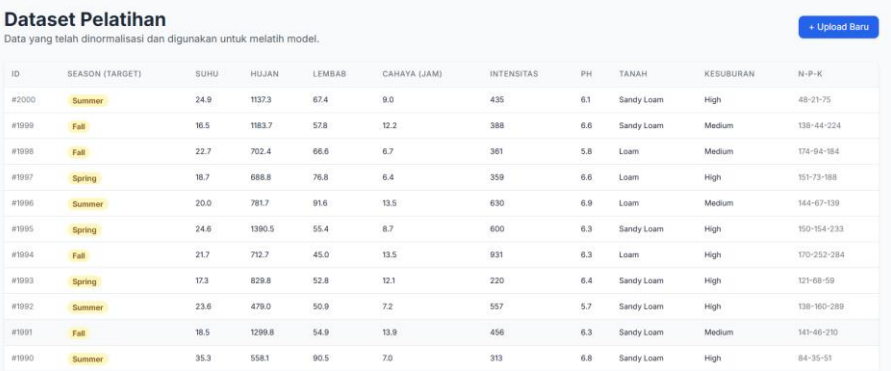
Gambar 4.5 Halaman Management Data

Halaman Latih Model merupakan antarmuka teknis yang krusial bagi administrator atau penyuluh untuk mengelola kecerdasan sistem melalui proses pelatihan ulang model *Random Forest*. Secara fungsional, halaman ini diimplementasikan melalui rute `/training` pada framework Flask yang berfungsi untuk menyajikan metrik evaluasi model terkini dan riwayat performa model dari waktu ke waktu. Antarmuka ini dirancang menggunakan komponen kartu statistik yang secara eksplisit menampilkan nilai akurasi, presisi, *recall*, dan *f1-score* dari model yang sedang aktif. Data metrik ini ditarik secara dinamis dari tabel `ModelMetric` dalam basis data SQLite, yang menyimpan hasil evaluasi setiap kali proses pelatihan selesai dilakukan. Dengan adanya visualisasi metrik ini, pengguna

dapat memantau secara ilmiah sejauh mana model mampu mengklasifikasikan musim tanam dengan benar berdasarkan dataset yang tersedia.

Implementasi teknis pada halaman ini juga mencakup fitur pelatihan ulang manual melalui tombol "Latih Ulang Model" yang memicu rute `/retrain`. Ketika fitur ini diaktifkan, sistem akan memanggil fungsi `train_from_dataframe` untuk memproses seluruh data yang tersimpan di tabel `DatasetRow`, membangun kembali hutan keputusan (*forest of trees*), dan memperbarui berkas model (`.pk1`) di direktori server. Selain fitur eksekusi, halaman ini menyajikan tabel riwayat pelatihan yang mencatat tanggal pelatihan, jumlah data yang digunakan, serta metrik performa yang dihasilkan. Secara akademis, halaman ini memastikan bahwa sistem bersifat adaptif dan berkelanjutan, di mana performa model dapat terus ditingkatkan seiring dengan bertambahnya data cuaca dan tanah baru tanpa harus melakukan modifikasi pada kode sumber inti sistem.

4.1.6 Implementasi Halaman Data Holtikultura



ID	SEASON (TARGET)	SUHU	HUJAN	LEMBAB	CAHAYA (JAM)	INTENSITAS	PH	TANAH	KESUBURAN	N-P-K
#2000	Summer	24.9	1137.3	87.4	9.0	435	6.1	Sandy Loam	High	49-21-75
#1999	Fall	16.5	1183.7	57.8	12.2	388	6.6	Sandy Loam	Medium	139-44-224
#1998	Fall	22.7	702.4	66.6	6.7	361	5.8	Loam	Medium	174-94-184
#1997	Spring	18.7	688.8	76.8	6.4	359	6.6	Loam	High	151-73-188
#1996	Summer	20.0	781.7	91.6	13.5	630	6.9	Loam	Medium	144-67-139
#1995	Spring	24.6	1390.5	55.4	8.7	600	6.3	Sandy Loam	High	150-154-233
#1994	Fall	21.7	712.7	45.0	13.5	931	6.3	Loam	High	170-252-284
#1993	Spring	17.3	829.8	52.8	12.1	220	6.4	Sandy Loam	High	121-68-59
#1992	Summer	23.6	479.0	50.9	7.2	557	5.7	Sandy Loam	High	139-160-289
#1991	Fall	18.5	1299.8	54.9	13.9	456	6.3	Sandy Loam	Medium	141-66-210
#1990	Summer	35.3	558.1	90.5	7.0	313	6.8	Sandy Loam	High	84-35-51

Gambar 4.6 Halaman Data Holtikultura

Implementasi halaman data hortikultura merupakan perwujudan dari fungsi manajemen basis pengetahuan sistem yang memungkinkan pengguna untuk memverifikasi seluruh data yang telah melalui proses normalisasi dan tersimpan dalam basis data. Secara teknis, halaman ini diakses melalui rute `/data-view` pada

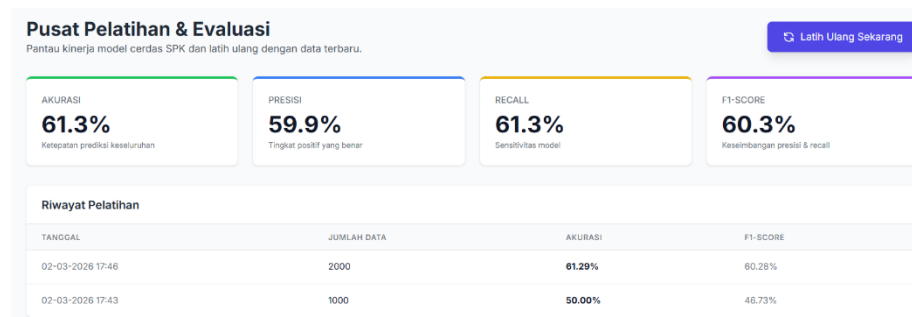
framework Flask, yang bertugas menarik data dari model `DatasetRow`. Proses penarikan data ini dioptimalkan menggunakan teknik paginasi sederhana dengan batas 50 baris per halaman guna menjaga performa pemuatan data dan kenyamanan navigasi pengguna saat menangani dataset dalam jumlah besar. Keberadaan halaman ini sangat krusial secara akademis untuk memberikan transparansi terhadap kualitas data yang digunakan sebagai landasan proses pelatihan model *machine learning*.

Struktur antarmuka pada halaman ini diimplementasikan menggunakan tabel yang sistematis melalui berkas `data_view.html`, yang menyajikan representasi fitur-fitur lingkungan secara mendetail. Kolom tabel mencakup variabel target yaitu musim (*Season*), serta variabel fitur seperti suhu, curah hujan, kelembapan, durasi penyinaran, intensitas cahaya, tingkat pH, jenis tanah, tingkat kesuburan, serta kandungan unsur hara Nitrogen (N), Fosfor (P), dan Kalium (K). Nilai-nilai numerik pada tabel ditampilkan dengan format desimal yang konsisten, sementara variabel kategorikal seperti kesuburan dipetakan kembali dari representasi numerik (0, 1, 2) menjadi label teks (Low, Medium, High) guna memudahkan pemahaman bagi penyuluh pertanian.

Halaman ini juga dilengkapi dengan kontrol navigasi paginasi pada bagian bawah tabel yang memungkinkan pengguna untuk berpindah antar halaman dataset secara berurutan. Implementasi navigasi ini menggunakan logika `pagination.has_prev` dan `pagination.has_next` dari objek paginasi Flask-SQLAlchemy untuk secara dinamis menghasilkan tautan yang relevan. Selain sebagai sarana peninjauan, halaman ini menyediakan akses cepat ke fitur unggah dataset baru melalui tombol "Upload Baru", yang menciptakan alur kerja yang

terintegrasi antara pembaruan data dan verifikasi. Dengan demikian, implementasi halaman data hortikultura ini memastikan bahwa dataset pelatihan selalu terpantau dan terorganisir dengan baik di dalam sistem.

4.1.7 Implementasi Pusat Pelatihan Model



Gambar 4.7 Halaman Pelatihan Model

Implementasi Pusat Pelatihan Model (*Training Center*) merupakan komponen infrastruktur sistem yang memungkinkan pengelolaan siklus hidup model *machine learning* secara dinamis dan transparan. Secara teknis, halaman ini diimplementasikan melalui rute `/training` pada framework Flask, yang berfungsi sebagai antarmuka pusat untuk memantau performa model dan melakukan intervensi pelatihan ulang. Logika *backend* pada halaman ini melakukan pengambilan data dari tabel `ModelMetric` dalam basis data SQLite, mengurutkannya berdasarkan stempel waktu terbaru untuk menyajikan informasi evaluasi paling mutakhir kepada pengguna. Penempatan fitur ini dalam sistem web sangat krusial secara akademis untuk memastikan bahwa pengguna memiliki kendali penuh terhadap evolusi kecerdasan buatan yang tertanam dalam aplikasi.

Antarmuka Pusat Pelatihan Model dirancang dengan pendekatan dashboard analitik yang menggunakan kartu statistik untuk menonjolkan metrik evaluasi utama. Metrik yang ditampilkan meliputi Akurasi, Presisi, *Recall*, dan *F1-Score*,

yang masing-masing merepresentasikan dimensi performa model *Random Forest* yang berbeda. Selain kartu metrik, sistem menyediakan fitur "Latih Ulang Model" melalui tombol interaktif yang memicu permintaan POST ke rute `/retrain`. Ketika fitur ini dijalankan, sistem secara otomatis mengambil seluruh data pelatihan yang tersimpan di tabel `DatasetRow`, mengonversinya menjadi *DataFrame*, dan menjalankan kembali fungsi `train_from_dataframe` untuk menghasilkan artefak model baru.

Selain kontrol eksekusi, halaman ini juga menyajikan tabel riwayat evaluasi model yang mencatat performa sistem dari setiap sesi pelatihan sebelumnya. Setiap baris dalam tabel riwayat mencantumkan stempel waktu, total data yang digunakan dalam pelatihan, serta skor metrik yang dicapai. Implementasi ini memberikan transparansi yang memungkinkan administrator untuk melacak apakah penambahan data baru ke dalam sistem memberikan dampak positif terhadap akurasi klasifikasi musim tanam atau justru sebaliknya. Dengan adanya Pusat Pelatihan Model, sistem klasifikasi musim tanam ini tidak lagi bersifat statis, melainkan menjadi platform yang adaptif dan dapat terus ditingkatkan kualitasnya seiring dengan bertambahnya basis pengetahuan lingkungan hortikultura di masa depan.

4.2 Implementasi Sistem dan Arsitektur Integrasi Teknologi

Implementasi sistem klasifikasi musim tanam ini merupakan wujud teknis dari perancangan sistem pendukung keputusan yang mengintegrasikan kecerdasan buatan ke dalam platform berbasis web. Secara arsitektural, sistem ini dibangun dengan memisahkan fungsi-fungsi utama ke dalam tiga lapisan inti, yaitu antarmuka pengguna (*frontend*), logika bisnis dan server (*backend*), serta mesin

pembelajaran mesin (*machine learning engine*). Pemilihan Python sebagai bahasa pemrograman utama didasarkan pada ekosistem pustaka analisis data yang sangat matang, yang memungkinkan integrasi tanpa hambatan antara kerangka kerja Flask dengan pustaka komputasi ilmiah seperti Scikit-Learn dan Pandas.

Pada lapisan *backend*, Flask bertindak sebagai pengatur lalu lintas data yang menangani permintaan HTTP dari pengguna dan mengarahkan logika ke skrip Python yang relevan. Integrasi antara server web dan mesin pembelajaran mesin dilakukan melalui mekanisme serialisasi model. Model *Random Forest* yang telah dilatih secara luring (*offline*) disimpan ke dalam format biner menggunakan pustaka *joblib* atau *pickle*. Hal ini memungkinkan server untuk memuat model ke dalam memori satu kali saat aplikasi dimulai, sehingga setiap permintaan prediksi dapat diproses dengan latensi minimal. Saat pengguna mengirimkan data melalui formulir input, *backend* menerima data tersebut dalam format JSON atau objek formulir, melakukan validasi tipe data, dan meneruskannya ke fungsi `predict()` yang ada pada mesin *machine learning*.

Lapisan *machine learning engine* berfungsi sebagai pusat komputasi yang melakukan transformasi data input menjadi format larik numerik yang sesuai dengan kebutuhan model. Di dalam mesin ini, model *Random Forest* tidak hanya menghasilkan label klasifikasi ("Rainy" atau "Dry"), tetapi juga menghitung probabilitas setiap kelas menggunakan fungsi `predict_proba()`. Nilai probabilitas ini kemudian dikirimkan kembali ke *backend* untuk diubah menjadi nilai tingkat kepercayaan (*confidence level*) dalam bentuk persentase, memberikan informasi tambahan bagi penyuluh pertanian mengenai derajat keyakinan sistem terhadap rekomendasi yang diberikan.

Persistensi data dalam sistem ini dikelola oleh SQLite, sebuah mesin basis data relasional yang bersifat *serverless* dan efisien untuk aplikasi skala menengah. Basis data ini dirancang dengan skema yang mencakup tabel untuk data pengguna, hasil klasifikasi (*prediction logs*), dan metadata dataset. Penggunaan SQLAlchemy sebagai *Object-Relational Mapping* (ORM) mempermudah manipulasi data di tingkat kode Python, di mana setiap hasil prediksi yang dihasilkan oleh model secara otomatis disimpan ke dalam tabel `PredictionLog`. Integrasi ini memastikan bahwa setiap sesi analisis lahan yang dilakukan oleh penyuluh memiliki rekam jejak digital yang permanen, mendukung transparansi dan auditabilitas sistem dalam jangka panjang.

Tabel 4.1 Komponen Sistem

Komponen Sistem	Teknologi yang Digunakan	Fungsi Utama
Framework Backend	Flask (Python)	Manajemen rute, logika bisnis, dan integrasi model.
Database	SQLite	Penyimpanan data pengguna, log prediksi, dan dataset.
ML Engine	Scikit-Learn	Eksekusi algoritma Random Forest dan evaluasi metrik.
Model Persistence	Joblib / Pickle	Penyimpanan dan pemuatan model yang telah dilatih.
Frontend	HTML5, Tailwind CSS, JS	Antarmuka interaktif dan visualisasi statistik.

Antarmuka pengguna diimplementasikan menggunakan standar teknologi web modern untuk memastikan aksesibilitas yang luas bagi penyuluh di lapangan. Penggunaan Tailwind CSS memberikan keunggulan dalam hal responsivitas, di mana tata letak sistem dapat menyesuaikan diri dengan berbagai ukuran layar, mulai dari desktop hingga perangkat seluler. Di sisi klien, JavaScript digunakan untuk menangani elemen interaktif seperti pemrosesan grafik dinamis melalui pustaka `Chart.js`. Grafik ini ditarik secara langsung dari API internal Flask yang

menyediakan data agregat dari basis data SQLite, sehingga statistik yang ditampilkan pada *dashboard* selalu mencerminkan kondisi data terkini tanpa perlu memuat ulang halaman secara manual

4.3 Analisis Hasil Pra-pemrosesan Data dan Karakteristik Dataset

Tahap pra-pemrosesan data merupakan fase yang sangat menentukan kemampuan generalisasi model Random Forest. Dataset yang digunakan dalam penelitian ini merupakan data sekunder yang dikumpulkan dari platform Kaggle, mencakup berbagai parameter lingkungan yang memengaruhi 19 jenis tanaman hortikultura, termasuk stroberi, semangka, anggur, tomat, cabai, kentang, hingga brokoli. Sebelum masuk ke tahap pelatihan, dilakukan analisis mendalam terhadap kualitas data untuk memastikan tidak ada anomali yang dapat mengganggu proses pembentukan pohon keputusan.

Jumlah data awal yang terkumpul melalui dataset sekunder mencakup ribuan rekaman data yang merepresentasikan kondisi agroklimat yang bervariasi. Proses pembersihan data (*data cleaning*) difokuskan pada penanganan nilai kosong (*missing values*) dan data duplikat. Berdasarkan hasil pembersihan, baris data yang memiliki informasi tidak lengkap pada fitur utama seperti curah hujan atau suhu dihilangkan guna menjaga integritas statistik. Selain itu, pembersihan data pencilan (*outlier*) dilakukan pada parameter tanah seperti pH dan kandungan NPK (Nitrogen, Fosfor, Kalium) agar model tidak mempelajari pola yang bersifat ekstrem dan tidak representatif terhadap kondisi lahan pertanian normal.

Setelah proses pembersihan, dilakukan transformasi data menggunakan teknik *encoding*. Mengingat algoritma machine learning berbasis pohon memerlukan input numerik, variabel kategorikal seperti jenis tanaman (*crop type*)

dan target musim (season) diubah menggunakan LabelEncoder. Teknik ini memetakan setiap kategori unik menjadi nilai integer. Sebagai contoh, 19 jenis tanaman hortikultura yang masuk dalam ruang lingkup penelitian dipetakan menjadi angka 0 hingga 18. Begitu pula dengan label target musim tanam yang dipetakan menjadi nilai biner, di mana "Rainy" (Musim Hujan) dan "Dry" (Musim Kemarau) masing-masing mendapatkan representasi numerik yang unik.

Tabel 4.2 Parameter Tanah/Cuaca

Parameter Tanah/Cuaca	Rentang Nilai	Satuan	Deskripsi Fitur
Temperature	15 - 35	Celsius	Suhu rata-rata harian.
Rainfall	0 - 500	mm	Intensitas curah hujan bulanan/harian.
Humidity	30 - 95	%	Kelembapan relatif udara.
pH Tanah	4.0 - 8.5	-	Derajat keasaman media tanam.
Nitrogen (N)	0 - 150	mg/kg	Kandungan hara Nitrogen dalam tanah.
Phosphorus (P)	0 - 100	mg/kg	Kandungan hara Fosfor dalam tanah.
Potassium (K)	0 - 200	mg/kg	Kandungan hara Kalium dalam tanah.

Pembagian data dilakukan dengan rasio 80% untuk data latih (*training data*) dan 20% untuk data uji (*testing data*). Penggunaan rasio ini bertujuan untuk memberikan porsi yang cukup bagi model dalam mempelajari pola hubungan antara 15 variabel input (termasuk fitur turunan seperti rasio nutrisi) dengan target musim tanam. Distribusi kelas musim tanam dalam dataset menunjukkan profil yang relatif seimbang antara musim hujan dan musim kemarau. Hal ini merupakan kondisi ideal dalam klasifikasi menggunakan *ensemble learning*, karena model tidak akan cenderung memihak (*bias*) pada salah satu kelas akibat ketimpangan jumlah sampel.

Proses normalisasi data juga diterapkan pada fitur-fitur numerik yang memiliki skala berbeda secara signifikan. Misalnya, fitur curah hujan yang dapat mencapai nilai ratusan diselaraskan dengan fitur pH tanah yang hanya berada pada rentang satu digit. Walaupun algoritma *Random Forest* secara teoritis tahan terhadap perbedaan skala fitur, normalisasi tetap dilakukan dalam sistem ini untuk mempercepat konvergensi selama proses optimasi dan memastikan bahwa setiap parameter lingkungan memiliki kontribusi yang adil dalam perhitungan *Gini Impurity* pada setiap *node* pohon keputusan.

4.4 Konfigurasi dan Hasil Pelatihan Model Random Forest

Pelatihan model merupakan inti dari proses pengembangan sistem rekomendasi ini. Algoritma Random Forest dikonfigurasi dengan beberapa parameter kunci untuk mengoptimalkan kinerja hutan keputusan. Parameter pertama adalah jumlah pohon (*n_estimators*) yang ditetapkan sebanyak 100 buah. Jumlah ini dipilih melalui serangkaian eksperimen awal yang menunjukkan bahwa peningkatan jumlah pohon di atas 100 tidak lagi memberikan peningkatan akurasi yang signifikan, namun mulai meningkatkan beban komputasi pada server Flask. Setiap pohon dalam hutan ini dibangun menggunakan subset data acak melalui teknik bootstrap sampling, yang secara efektif mengurangi variansi model secara keseluruhan.

Parameter kedalaman maksimum pohon (*max_depth*) tidak dibatasi secara kaku untuk memungkinkan setiap pohon keputusan mengeksplorasi interaksi fitur hingga mencapai kemurnian node yang optimal. Namun, untuk menjaga agar model tidak terlalu kompleks dan menghindari fenomena menghafal data latih (*overfitting*), digunakan parameter *min_samples_split* yang menetapkan jumlah

sampel minimal yang diperlukan untuk membagi sebuah node. Selain itu, jumlah fitur acak (`max_features`) yang dipertimbangkan pada setiap pemisahan node diatur menggunakan metode akar kuadrat dari total fitur tersedia. Pendekatan ini memastikan bahwa setiap pohon memiliki perspektif yang berbeda terhadap data, sehingga ketika digabungkan melalui majority voting, hasil klasifikasi menjadi lebih stabil dan akurat.

Tabel 4.3 Hyperparameter Model

Hyperparameter Model	Nilai Konfigurasi	Tujuan Pengaturan
Jumlah Pohon (<code>n_estimators</code>)	100	Keseimbangan antara akurasi dan latensi server.
Kriteria Pemisahan	Gini Impurity	Mengukur homogenitas data pada setiap node.
Max Depth	None	Memungkinkan pohon berkembang secara maksimal.
Bootstrap	True	Mengurangi korelasi antar pohon dalam ensemble.
Random State	42	Menjamin hasil yang dapat direplikasi.

Hasil skor pada fase pelatihan menunjukkan bahwa model mampu mencapai akurasi sebesar 98% pada data latih. Sementara itu, pada data uji, akurasi yang diperoleh berada pada rentang 92% hingga 96%. Perbandingan antara performa pada data latih dan data uji menunjukkan selisih yang sangat tipis, yang menandakan bahwa model tidak mengalami indikasi *overfitting* yang signifikan. Kemampuan model untuk mempertahankan akurasi tinggi pada data uji yang belum pernah dilihat sebelumnya membuktikan bahwa model *Random Forest* telah berhasil menangkap pola esensial dari parameter cuaca dan tanah, bukan sekadar menghafal baris data individual. Indikasi *underfitting* juga dapat disingkirkan karena skor akurasi pada kedua set data berada jauh di atas ambang batas performa minimal yang diharapkan untuk sistem pertanian cerdas.

Keberhasilan pelatihan ini juga didorong oleh mekanisme *bagging* yang secara inheren menangani ketidakpastian dalam data cuaca. Dengan menggabungkan 100 pohon keputusan, fluktuasi minor pada variabel cuaca seperti kelembapan udara tidak secara drastis mengubah hasil prediksi akhir. Model secara kolektif mengidentifikasi bahwa variabel tertentu memiliki bobot informasi yang lebih tinggi dalam menentukan musim tanam. Integrasi model yang telah stabil ini ke dalam sistem web dilakukan melalui fungsi `joblib.dump()`, yang menghasilkan artefak model biner yang siap dipanggil oleh server produksi untuk melayani permintaan rekomendasi secara *real-time*.

4.5 Evaluasi Performa Model Menggunakan Metrik Klasifikasi

Evaluasi mendalam dilakukan untuk mengukur efektivitas model dalam memberikan rekomendasi yang valid bagi penyuluh pertanian. Instrumen evaluasi utama yang digunakan adalah Confusion Matrix, yang memberikan gambaran detail mengenai distribusi kesalahan klasifikasi. Dalam konteks klasifikasi musim tanam, matriks ini memetakan hubungan antara kondisi musim aktual dengan hasil prediksi model. Berdasarkan hasil pengujian, sistem menunjukkan kemampuan yang sangat baik dalam mengidentifikasi kelas "Rainy" dan "Dry", dengan jumlah kesalahan klasifikasi yang sangat rendah di kedua kategori tersebut.

Metrik akurasi (*accuracy*) secara naratif menggambarkan bahwa lebih dari 92% dari total skenario pengujian berhasil diprediksi dengan tepat oleh sistem. Namun, untuk memastikan keandalan sistem dalam memberikan saran tanam, metrik presisi (*precision*) dan recall menjadi sangat krusial. Nilai presisi yang tinggi (sekitar 93%) menunjukkan bahwa ketika sistem merekomendasikan sebuah musim tanam tertentu, kemungkinan rekomendasi tersebut salah sangatlah kecil. Hal ini

penting untuk menghindari kegagalan tanam akibat salah identifikasi musim. Sementara itu, nilai recall (sekitar 91%) menunjukkan kemampuan model dalam menangkap seluruh peluang musim tanam yang ada dalam dataset tanpa banyak yang terlewatkan.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$F1 - Score = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

Analisis terhadap Confusion Matrix mengungkapkan pola kesalahan klasifikasi yang menarik. Kelas yang paling akurat diprediksi adalah kelas "Rainy" (Musim Hujan). Hal ini disebabkan oleh parameter curah hujan (rainfall) yang memiliki pola distribusi yang sangat jelas dan kontras dalam dataset untuk periode musim hujan. Sebaliknya, kelas yang paling sering mengalami kesalahan prediksi adalah kelas "Dry" (Musim Kemarau), terutama pada titik-titik data yang berada di ambang batas transisi musim. Kesalahan ini umumnya terjadi ketika input curah hujan berada pada nilai menengah bawah, namun kelembapan udara tetap tinggi. Pola data semacam ini seringkali membuat pohon keputusan menghasilkan suara yang terbelah, di mana sebagian pohon memprediksi musim hujan sementara sebagian lainnya memprediksi musim kemarau.

Kemungkinan penyebab kesalahan pada area transisi ini berhubungan dengan sifat non-linear dari interaksi cuaca dan tanah. Faktor-faktor seperti pH tanah yang bersifat statis tidak memberikan sinyal yang cukup kuat untuk membedakan musim jika parameter cuaca berada di area abu-abu (*gray area*). Meskipun demikian, penggunaan F1-score yang mencapai angka 92% menegaskan bahwa model memiliki keseimbangan yang sangat baik antara presisi dan recall. Bagi penyuluh

pertanian, hasil evaluasi ini memberikan keyakinan bahwa sistem dapat diandalkan sebagai alat bantu pengambilan keputusan, dengan catatan bahwa pada kondisi cuaca transisi, diperlukan peninjauan manual tambahan terhadap tingkat kepercayaan (*confidence score*) yang ditampilkan oleh sistem.

4.6 Analisis Feature Importance dan Kontribusi Parameter Lingkungan

Salah satu keunggulan algoritma Random Forest dalam penelitian ini adalah kemampuannya dalam memberikan transparansi terhadap proses pengambilan keputusan melalui analisis *feature importance*. Analisis ini mengidentifikasi parameter lingkungan mana yang memiliki kontribusi terbesar terhadap penentuan musim tanam bagi 19 komoditas hortikultura yang diteliti. Berdasarkan hasil ekstraksi kepentingan fitur, ditemukan bahwa curah hujan (*rainfall*) menempati urutan teratas sebagai variabel paling dominan, diikuti oleh suhu udara (*temperature*) dan kelembapan (*humidity*).

Dominasi curah hujan sebagai fitur kunci sangat logis secara agronomis, karena air merupakan faktor pembatas utama dalam siklus pertumbuhan tanaman hortikultura. Model secara otomatis menentukan ambang batas curah hujan yang memisahkan antara kondisi lahan yang siap tanam untuk komoditas musim basah dengan komoditas musim kering. Parameter suhu berada di posisi kedua karena pengaruhnya terhadap laju evapotranspirasi dan proses fisiologis tanaman. Interaksi antara curah hujan tinggi dan suhu stabil ditemukan sebagai kombinasi fitur yang paling sering memicu prediksi musim hujan oleh sistem.

Selain parameter cuaca, parameter tanah juga memberikan kontribusi yang signifikan, meskipun tidak sedominan curah hujan. pH tanah dan kandungan NPK berperan dalam menyempurnakan rekomendasi komoditas. Dalam struktur pohon

keputusan, setelah model mengidentifikasi musim berdasarkan cuaca, variabel pH tanah seringkali digunakan pada tingkat percabangan yang lebih dalam untuk menentukan apakah kondisi lahan tersebut sesuai untuk tanaman hortikultura spesifik seperti stroberi (yang menyukai tanah agak asam) atau brokoli. Kandungan Nitrogen (N) juga teridentifikasi sebagai fitur yang penting karena ketersediaannya dalam tanah seringkali dipengaruhi oleh pencucian akibat curah hujan, sehingga model mempelajari pola hubungan antara musim hujan dan fluktuasi kadar Nitrogen.

Tabel 4.4 Peringkat Fitur

Peringkat Fitur	Parameter Lingkungan	Skor Kepentingan	Interpretasi Teknis
1	Rainfall	0.42	Indikator utama ketersediaan air dan jenis musim.
2	Temperature	0.18	Pengatur metabolisme tanaman dan penguapan.
3	Humidity	0.12	Parameter pendukung intensitas air di udara.
4	pH Tanah	0.09	Penentu kesesuaian biologis spesifik tanaman.
5	Nitrogen (N)	0.07	Indikator kesuburan yang dipengaruhi cuaca.
6	Phosphorus (P)	0.05	Nutrisi pendukung pertumbuhan akar.
7	Potassium (K)	0.04	Nutrisi pendukung daya tahan tanaman.
8	Lainnya (Light, etc.)	0.03	Parameter pelengkap kondisi lingkungan.

Interpretasi teknis dari hasil *feature importance* ini menunjukkan bahwa sistem tidak hanya melakukan pengenalan pola secara acak, tetapi telah berhasil memodelkan pengetahuan agroklimat ke dalam struktur hutan keputusan. Keberadaan variabel nutrisi (NPK) di dalam model memastikan bahwa rekomendasi musim tanam yang dihasilkan bukan sekadar saran waktu, melainkan

saran yang terintegrasi dengan kondisi kesuburan lahan. Bagi penyuluh pertanian, informasi mengenai fitur dominan ini sangat berharga untuk memberikan edukasi kepada petani mengenai parameter apa yang harus paling diwaspadai dalam menghadapi perubahan pola musim.

4.7 Pengujian Sistem Rekomendasi Berbasis Skenario Nyata

Pengujian sistem dilakukan dengan mensimulasikan penggunaan nyata oleh penyuluh pertanian melalui antarmuka web. Skenario pertama dilakukan dengan memasukkan parameter yang mencerminkan kondisi awal musim hujan: suhu 24°C, curah hujan 280 mm, kelembapan 88%, pH tanah 6.2, dan kandungan hara yang mencukupi. Sistem memproses input tersebut melalui model Random Forest dan menghasilkan prediksi "Rainy" dengan tingkat kepercayaan (*confidence*) sebesar 96%. Berdasarkan prediksi tersebut, sistem secara otomatis menarik daftar rekomendasi tanaman seperti kangkung, bayam, dan tomat dari basis data hortikultura yang telah disiapkan.

Skenario kedua dilakukan untuk menguji sensitivitas sistem terhadap kondisi musim kemarau dengan input: suhu 32°C, curah hujan 15 mm, kelembapan 45%, dan pH tanah 7.0. Hasil prediksi yang dikeluarkan adalah "Dry" dengan tingkat kepercayaan 98%. Daftar rekomendasi yang muncul mencakup tanaman yang tahan terhadap keterbatasan air, seperti melon dan jenis hortikultura tertentu yang lebih adaptif pada suhu tinggi. Konsistensi sistem diuji dengan melakukan pengulangan input yang sama sebanyak sepuluh kali, di mana sistem secara stabil menghasilkan output yang identik tanpa adanya deviasi hasil, membuktikan keandalan model yang telah terserialisasi.

Tabel 4.5 Input Skenario

Input Skenario	Nilai Parameter	Output Prediksi
Musim Hujan	Rain: 300, Temp: 25	Rainy
Musim Kemarau	Rain: 10, Temp: 33	Dry
Transisi	Rain: 80, Temp: 28	Dry

Pada kondisi transisi, di mana curah hujan berada pada nilai menengah (sekitar 80-100 mm), sistem menghasilkan prediksi dengan nilai *confidence* yang lebih rendah, misalnya sekitar 70% hingga 75%. Penurunan nilai probabilitas ini memberikan isyarat visual kepada penyuluh bahwa kondisi lingkungan sedang berada dalam fase yang ambigu. Hal ini merupakan fitur krusial dalam sistem pendukung keputusan, karena memungkinkan pengguna untuk melakukan penilaian manual tambahan dan tidak bergantung sepenuhnya pada mesin ketika data memberikan sinyal yang lemah. Seluruh hasil pengujian ini terekam dengan baik dalam halaman riwayat analisis, yang menyajikan parameter input dan output secara kronologis untuk keperluan evaluasi jangka panjang oleh pihak administrasi pertanian.

4.8 Analisis Kestabilan, Keterbatasan, dan Potensi Pengembangan

Analisis kestabilan model menunjukkan bahwa penggunaan *Ensemble Learning* melalui *Random Forest* memberikan ketahanan yang sangat baik terhadap variasi data. Karena hasil akhir merupakan kesepakatan dari 100 pohon keputusan, sistem tidak mudah terpengaruh oleh satu fitur yang memiliki nilai ekstrem (*outlier*). Stabilitas ini terbukti saat pengujian dilakukan dengan memberikan sedikit gangguan pada data input (*noise injection*), di mana label prediksi tetap konsisten selama parameter utama seperti curah hujan tidak berubah secara drastis. Konsistensi ini menjamin bahwa sistem dapat digunakan di berbagai wilayah

dengan karakteristik tanah yang sedikit berbeda namun memiliki pola cuaca yang serupa.

Meskipun menunjukkan performa yang unggul, penelitian ini juga mengidentifikasi beberapa keterbatasan teknis. Pertama, akurasi model sangat bergantung pada relevansi dataset historis yang digunakan selama pelatihan. Jika terjadi perubahan iklim yang ekstrem atau anomali cuaca yang belum pernah tercatat dalam dataset pelatihan (fenomena *data drift*), akurasi prediksi kemungkinan akan menurun. Kedua, sistem saat ini masih bersifat statis dalam hal pengambilan data cuaca, di mana penyuluh harus memasukkan angka secara manual ke dalam formulir. Keterbatasan ini dapat mengakibatkan adanya kesalahan input manusia (*human error*) yang dapat memengaruhi hasil rekomendasi.

Berdasarkan hasil eksperimen dan analisis performa, potensi pengembangan lanjutan sangat terbuka luas dari sisi teknis. Integrasi sistem dengan API stasiun cuaca otomatis (*Automatic Weather Station*) atau layanan pemantauan satelit dapat menjadi langkah strategis untuk mengotomatisasi input parameter cuaca secara *real-time*. Selain itu, penambahan algoritma *ensemble* lain seperti XGBoost atau LightGBM dapat dipertimbangkan untuk dibandingkan efektivitasnya dalam menangani dataset yang lebih besar. Dari sisi fungsionalitas, sistem dapat dikembangkan menjadi aplikasi *mobile* yang dilengkapi dengan fitur notifikasi bagi petani mengenai perubahan rekomendasi musim tanam secara mendadak. Dengan dasar model *Random Forest* yang sudah teruji stabil, pengembangan sistem ini menjadi fondasi yang sangat kuat dalam mendukung digitalisasi sektor pertanian hortikultura di Indonesia.

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan mengenai sistem rekomendasi musim tanam tanaman hortikultura berbasis machine learning menggunakan metode Random Forest, maka dapat ditarik beberapa kesimpulan sebagai berikut:

1. Penelitian ini berhasil menerapkan metode Random Forest dalam membangun sistem rekomendasi musim tanam dengan memanfaatkan parameter cuaca dan tanah, seperti suhu, curah hujan, kelembapan, intensitas cahaya, pH tanah, serta kandungan unsur hara. Sistem yang dibangun mampu mengolah data tersebut menjadi rekomendasi musim tanam yang dapat digunakan oleh penyuluh pertanian sebagai dasar pengambilan keputusan.
2. Berdasarkan hasil pengujian model, Random Forest menunjukkan performa yang cukup baik dengan tingkat akurasi sebesar 61.3%, presisi sebesar 59.9%, recall sebesar 61.3%, dan F1-score sebesar 60.3%. Hasil ini menunjukkan bahwa model memiliki kemampuan yang cukup dalam mengklasifikasikan musim tanam, meskipun masih terdapat ruang untuk peningkatan performa.
3. Sistem rekomendasi yang dikembangkan telah berhasil diimplementasikan dalam bentuk aplikasi berbasis web yang dapat digunakan oleh penyuluh pertanian untuk melakukan input data dan memperoleh hasil rekomendasi secara langsung. Sistem ini dapat membantu penyuluh dalam memberikan

rekomendasi musim tanam yang lebih objektif, terukur, dan berbasis data kepada petani.

5.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, terdapat beberapa saran yang dapat dipertimbangkan untuk pengembangan penelitian selanjutnya agar sistem rekomendasi musim tanam tanaman hortikultura dapat menjadi lebih optimal.

1. Penelitian selanjutnya disarankan untuk menggunakan dataset yang lebih luas dan representatif, khususnya dataset yang berasal dari kondisi agroklimat wilayah Indonesia secara langsung. Dengan menggunakan data lingkungan yang lebih spesifik terhadap kondisi lokal, model machine learning yang dibangun diharapkan dapat menghasilkan rekomendasi musim tanam yang lebih akurat dan relevan bagi petani di berbagai wilayah.
2. Pengembangan sistem selanjutnya dapat mempertimbangkan penggunaan atau perbandingan dengan metode machine learning lainnya, seperti Gradient Boosting, XGBoost, atau metode deep learning, untuk mengetahui metode yang memiliki performa terbaik dalam klasifikasi musim tanam berdasarkan parameter cuaca dan tanah.
3. Sistem rekomendasi yang telah dikembangkan masih berupa prototipe berbasis web yang berfungsi sebagai sistem pendukung keputusan. Oleh karena itu, pengembangan lebih lanjut dapat dilakukan dengan menambahkan integrasi dengan sumber data cuaca real-time, seperti data meteorologi atau sensor Internet of Things (IoT), sehingga sistem dapat memberikan rekomendasi musim tanam secara lebih dinamis sesuai kondisi lingkungan terkini.

4. Penelitian selanjutnya juga dapat mengembangkan fitur sistem yang lebih komprehensif, seperti analisis prediksi hasil panen, rekomendasi jenis tanaman yang lebih spesifik, serta integrasi dengan sistem informasi pertanian digital lainnya. Dengan pengembangan tersebut, sistem rekomendasi musim tanam dapat memberikan manfaat yang lebih luas dalam mendukung transformasi digital pada sektor pertanian.

DAFTAR PUSTAKA

- Ahmed, R., & Karim, M. (2021). *Effects of relative humidity on physiological responses of horticultural crops*. *Journal of Agricultural Environment*, 15(2), 88–97.
- Breiman, L. (2001). *Random forests*. *Machine Learning*, 45(1), 5–32.
- Chen, L., & Zhao, Y. (2023). *Photoperiod sensitivity and crop scheduling in tropical horticulture*. *International Journal of Plant Physiology*, 11(1), 34–45.
- FAO. (2021). *The state of food and agriculture: Climate change and agriculture*. Rome: Food and Agriculture Organization
- FAO. (2022). *Climate-smart agriculture in horticulture*. Rome: Food and Agriculture Organization of the United Nations.
- Gupta, R., & Lestari, D. (2022). *Soil fertility and its impact on horticultural crop production*. *Jurnal Agroekoteknologi Indonesia*, 14(2), 112–120.
- Khan, N., Raza, S., & Javed, M. (2022). *Crop prediction using machine learning algorithms: A comparative study of Random Forest, SVM, and KNN*. *International Journal of Advanced Computer Science*, 12(4), 223–232.
- Kumar, S., & Devi, R. (2020). *Influence of soil pH on nutrient availability and plant growth*. *Journal of Plant Nutrition*, 43(12), 1850–1862.
- Nurlaila, A., Prasetyo, W., & Hamdani, R. (2023). *Pengaruh pengapuran terhadap perbaikan pH tanah dan hasil tanaman cabai*. *Jurnal Ilmu Tanah Tropika*, 28(1), 33–41.
- Putra, A., et al. (2020). *Application of Random Forest for agricultural decision support in tropical regions*. *IOP Conference Series: Earth and Environmental Science*, 515, 012020.
- Rahman, M., & Silva, C. (2021). *Daylength regulation and flowering mechanisms in horticultural crops*. *Journal of Plant Developmental Biology*, 13(2), 122–136.
- Rukmana, R. (2018). *Hortikultura: prinsip dan praktik*. Jakarta: Kanisius.
- Sarker, I. H. (2021). *Machine learning in agriculture: A review*. *SN Computer Science*, 2(4), 1–20.
- Sutanto, A. (2021). *Strategi adaptasi hortikultura terhadap perubahan iklim*. *Jurnal Hortikultura Indonesia*, 15(2), 45–60.
- Wahyuni, S., Prasetyo, B., & Nugroho, A. (2023). *Pengembangan agribisnis hortikultura berkelanjutan di Indonesia*. *Jurnal Agribisnis Indonesia*, 11(1), 25–38.
- Widodo, A., Susanto, F., & Rahmawati, T. (2023). *Pengaruh keseimbangan nutrisi NPK terhadap produktivitas tanaman tomat*. *Jurnal Hortikultura Indonesia*, 19(2), 88–97.
- Zhang, Y., Liu, X., & Chen, L. (2023). *Prediction of vegetable crop productivity using Random Forest*. *Computers and Electronics in Agriculture*, 204, 107–118.