



Penerapan Random Forest untuk Klasifikasi Penyakit Tanaman Padi

Irna Liana Hulu^{1*}, Nina Mumtazia², Siti Shapira Nurulita³, Ayu Widya Rakhmawati⁴

^{1,2,3,4}Fakultas Ilmu Komputer, Teknik Informatika, Universitas Pamulang, Kota Tangerang Selatan, Indonesia

Email: ^{1*}irnalianahulu@gmail.com, ²ninamumtazia47@gmail.com, ³sitishapiranurulita@gmail.com,

⁴ayuw2812@gmail.com

(* : coressponding author)

Abstrak – Literatur review ini mengkaji penerapan algoritma Random Forest dalam klasifikasi penyakit pada tanaman padi, salah satu komoditas utama dalam sektor pertanian Indonesia. Algoritma Random Forest dipilih karena kemampuannya dalam menangani dataset yang kompleks dan beragam, serta akurasi yang tinggi dalam berbagai skenario klasifikasi. Literatur review ini membahas studi-studi sebelumnya yang mengimplementasikan Random Forest dan membandingkannya dengan algoritma lain seperti SVM dan Decision Tree. Meskipun memiliki keunggulan dalam akurasi, Random Forest menghadapi tantangan dalam waktu komputasi pada dataset besar dan ketergantungan pada kualitas data. Hasil penelitian ini diharapkan menjadi referensi dalam pengembangan aplikasi pertanian berbasis machine learning untuk membantu petani dalam deteksi dan pengendalian penyakit tanaman padi secara efisien.

Kata Kunci: Random Forest, Klasifikasi, Penyakit Tanaman Padi, Machine Learning, Literatur Review

Abstract – This literature review examines the application of the Random Forest algorithm in disease classification in rice plants, one of the main commodities in the Indonesian agricultural sector. The Random Forest algorithm was chosen because of its ability to handle complex and diverse datasets, as well as its high accuracy in various classification scenarios. This literature review discusses previous studies that implement Random Forest and compares it with other algorithms such as SVM and Decision Tree. Despite its advantages in accuracy, Random Forest faces challenges in computing time on large data sets and dependence on data quality. It is hoped that the results of this research will become a reference in developing machine learning-based agricultural applications to help farmers detect and control rice plant diseases efficiently.

Keywords: Random Forest, Classification, Rice Plant Diseases, Machine Learning, Literature Review

1. PENDAHULUAN

Dalam klasifikasi suatu data, salah satu metode yang dapat digunakan adalah metode Decision Tree yang merupakan salah satu metode pengambilan keputusan. Pertumbuhan pada metode Decision Tree dapat ditumbuhkan sampai mencapai kehomogenan, tetapi hal tersebut dapat mengakibatkan overfitting, sehingga diperlukan suatu metode yang bisa mengatasi masalah tersebut. Metode yang dapat digunakan untuk menghindari overfitting adalah metode dengan memangkas pohon (pruning) atau menggunakan metode Random Forest (RF) (Amaliah, S., Nusrang, M., & Aswi., 2022).

Random forest merupakan salah satu algoritma machine learning yang menggabungkan keunggulan dari beberapa pohon keputusan untuk meningkatkan akurasi prediksi. Dalam pendekatan ini, kumpulan pohon keputusan dibuat, masing masing dilatih pada subset acak dari data set dan menggunakan sebagian kecil fitur. Setiap pohon membuat prediksi secara independen dan klasifikasi akhir ditentukan dengan menggabungkan hasil keluaran mereka melalui mekanisme voting atau perataan rata rata (Amaliah, S., Nusrang, M., & Aswi., 2022).

Pertanian sebagai salah satu sektor industri menjadi bagian pekerjaan yang menunjang pemenuhan kebutuhan makanan pokok masyarakat seperti tanaman pangan. Tanaman padi merupakan tanaman pangan yang rentan terserang hama. Pengenalan terhadap jenis hama yang menyerang merupakan langkah awal yang sangat penting untuk menunjang keberhasilan dalam usaha pengendaliannya. Hama tanaman padi tersebut dapat menjadi kendala bagi petani untuk bisa meningkatkan produksi. karena hama tersebut dapat merusak tanaman padi hingga membuat gagal panen. Oleh sebab itu perlu dilakukan deteksi klasifikasi pada hama daun padi untuk mencari akurasi dengan menggunakan perbandingan berbagai macam metode algoritma yaitu dengan Decision Tree,



Random Forest, Naïve Bayes, SVM dan KNN. Sehingga diharapkan mampu menangani hama secara tepat, agar tidak terjadi kerusakan dan gagal panen. Dengan menggunakan dataset Rice Leaf Diseases Detection untuk Deteksi dan Klasifikasi Penyakit Padi (Purnamawati, A., Nugroho, W., Putri, D., & Hidayat, W. F., 2020). Pertanian padi memiliki peran yang sangat penting dalam mendukung ketahanan pangan. Penyakit yang ada di tanaman padi menjadi salah satu tantangan utama dalam meningkatkan produktivitas dan kualitas pada panen padi. Pada penyakit tanaman padi dapat mungurangi kinerja dan hasil panen padi dan berpengaruh besar pada petani padi dan juga bisa mempengaruhi kualitas biji padi, dan secara umum akan sangat merugikan masyarakat petani padi. Berdasarkan penelitian ini bertujuan untuk mengembangkan pada metode klasifikasi yang efektif dalam mengidentifikasi penyakit yang ada di tanaman padi, dengan memanfaatkan suatu teknologi Random Forest, dengan tujuan berdasarkan penelitian untuk mengklasifikasikan penyakit pada tanaman padi (Rozi, K., Muhsi, Anwari, & Tamam, M. B., 2024).

2. METODOLOGI PENELITIAN

2.1 Systematic Literatur Review (SLR)

SLR atau dalam bahasa indonesia disebut tinjauan pustaka sistematis adalah metode *literature review* yang mengidentifikasi, menilai, dan menginterpretasi seluruh temuan-temuan pada suatu topik penelitian, untuk menjawab pertanyaan penelitian (*research question*) yang telah ditetapkan sebelumnya (Kitchenham & Charters, 2007). Systematic Literature Review (SLR) menganalisis berbagai penelitian mengenai topik tersebut. Data dikumpulkan dari jurnal ilmiah, laporan penelitian, dan publikasi terkait lainnya yang relevan Hasil Literatur telah menjawab *Research Question* (Maulana, A., & Sari, N. L. D. M., 2024).

2.2 Random Forest

Random forest didefinisikan sebagai teknik ensemble learning yang menggabungkan beberapa pohon keputusan untuk meningkatkan akurasi dan mengurangi risiko overfitting. Mereka menekankan bahwa random forest efektif dalam klasifikasi karena menggunakan metode voting untuk mendapatkan hasil akhir (Sharma dan Jain.,2020).

2.3 Klasifikasi

Pengertian klasifikasi adalah pengelompokan yang berarti mengumpulkan benda yang sama dan memisahkan benda yang tidak sama (Gatot Subrata.,2022). Penelitian ini melalui tahapan augmentasi data untuk meningkatkan variasi citra sebelum melakukan klasifikasi dengan random forest dan mendapatkan hasil akurasi yang cukup tinggi.

2.4 Research Question

Research question penting dalam studi sosial, di mana pertanyaan penelitian harus mampu menangkap kompleksitas dari masalah sosial yang diteliti. Pertanyaan penelitian yang baik dapat membantu peneliti lebih fokus pada isu-isu utama dan menghindari pengumpulan data yang tidak relevan (Bryman., 2015).

Tabel 1. Pertanyaan Peneliti

ID	Pertanyaan Peneliti
RQ1	Bagaimana algoritma Random Forest diterapkan dalam klasifikasi penyakit tanaman padi?
RQ2	Seberapa efektif Random Forest dibandingkan dengan algoritma lain dalam konteks klasifikasi penyakit tanaman padi?
RQ3	Apa tantangan dan solusi yang ditemukan dalam penelitian sebelumnya terkait penggunaan Random Forest pada klasifikasi penyakit tanaman padi?



2.5 Search Process

Pada tahap ini, penulis melakukan pencarian yang terstruktur untuk mengumpulkan data yang relevan dengan topik penelitian. Makalah atau literatur yang berkaitan dengan masalah penelitian diperoleh melalui sumber informasi seperti jurnal. Penulis memulai pencarian dengan menggunakan fungsi pencarian di Google Chrome atau aplikasi Publish or Perish untuk menemukan referensi yang dibutuhkan, kemudian mengunjungi situs <https://scholar.google.com>. Penulis memilih kata kunci yang relevan dengan topik penelitian. Untuk penelitian ini, digunakan kata kunci "Random forest " dan "penerapan random forest dalam klasifikasi penyakit tanaman padi ". Setelah menentukan rentang tahun publikasi antara 2013 hingga 2023, kata kunci tersebut dimasukkan dalam pencarian Google Scholar maupun tempat pencarian lainnya

2.6 Inclusion and Exclusion Criteria

Penulis mengevaluasi artikel-artikel yang ditemukan selama proses pencarian, dengan memperhatikan kriteria inklusi dan eksklusi yang telah ditetapkan sebelumnya.

Inclusion Criteria

a. Jenis Studi

Studi yang menggunakan algoritma Random Forest untuk klasifikasi penyakit tanaman, khususnya tanaman padi. Studi yang mengimplementasikan metode pembelajaran mesin dalam bidang pertanian, terutama untuk deteksi penyakit tanaman. Artikel yang berupa penelitian asli, baik eksperimen lapangan, model komputer, atau kajian literatur.

b. Populasi

Penelitian yang berfokus pada tanaman padi (*Oryza sativa*) sebagai objek penelitian utama. Penyakit tanaman padi yang dijadikan objek klasifikasi, baik itu penyakit bakteri, jamur, atau virus.

c. Batasan Waktu

Studi yang diterbitkan dalam 10 tahun terakhir (2013–2023), untuk mencakup perkembangan terbaru dalam teknologi Random Forest dan aplikasi dalam pertanian.

d. Bahasa

Artikel yang diterbitkan dalam bahasa Inggris atau bahasa Indonesia.

e. Metodologi

Penelitian yang mengkaji atau menggunakan algoritma Random Forest dalam konteks klasifikasi atau deteksi penyakit tanaman. Studi yang memberikan hasil evaluasi atau akurasi dalam penggunaan Random Forest untuk klasifikasi penyakit tanaman.

Exclusion Criteria

a. Jenis Studi

Studi yang tidak mengaplikasikan algoritma Random Forest (misalnya, menggunakan algoritma lain seperti SVM atau Neural Networks). Studi yang berupa opini, ulasan, atau artikel yang tidak menyediakan data eksperimen yang relevan. Studi yang tidak mengkhususkan pada klasifikasi penyakit tanaman padi.

b. Populasi

Studi yang berfokus pada tanaman selain padi, seperti tanaman jagung, tomat, atau tanaman lainnya. Penelitian yang tidak melibatkan penyakit tanaman atau hanya berfokus pada karakteristik tanaman tanpa klasifikasi penyakit.

c. Batasan Waktu

Artikel yang diterbitkan lebih dari 10 tahun yang lalu (sebelum 2013), untuk menghindari referensi yang tidak relevan dengan perkembangan terbaru.



d. Bahasa

Artikel yang ditulis dalam bahasa selain bahasa Inggris atau bahasa Indonesia (misalnya, artikel dalam bahasa yang tidak dapat dipahami atau diterjemahkan dengan mudah).

e. Metodologi

Penelitian yang tidak menyediakan hasil evaluasi atau kinerja algoritma Random Forest dalam konteks klasifikasi penyakit (misalnya, penelitian yang hanya menjelaskan dasar teori tanpa aplikasi praktis atau hasil pengujian).

2.7 Data Collection

Menyatakan bahwa pengumpulan data mencakup kegiatan untuk memperoleh data yang sesuai dengan objek penelitian, baik secara langsung (primer) maupun dari sumber lain yang relevan (sekunder). Proses ini perlu dirancang dengan baik agar data yang diperoleh tidak hanya akurat tetapi juga representatif terhadap populasi yang diteliti (Nazir.,2020).

2.7.1 Data Primer

Data primer adalah data yang dikumpulkan secara langsung dari lapangan oleh peneliti sendiri, yang mencakup metode-metode seperti kuesioner atau observasi, untuk memastikan data yang relevan dengan studi yang dilakukan (Kuncoro.,2013). Data ini disusun berdasarkan karakteristik umum dari penyakit pada daun padi yang sering dijadikan dasar dalam penelitian, termasuk atribut visual dan kondisi lingkungan.

a. Deskripsi Dataset

Dataset ini terdiri dari beberapa fitur yang menggambarkan gejala penyakit pada daun padi, berdasarkan citra atau pengamatan langsung. Atribut yang sering digunakan dalam klasifikasi penyakit padi meliputi:

- Warna Daun: Hijau, Kuning, Coklat
- Jumlah dan Bentuk Bercak: Ada/Tidak Ada, Kecil/Sedang/Besar, Bulat/Tidak Beraturan
- Tekstur Daun: Kasar/Halus
- Kondisi Batang: Normal atau Busuk
- Kelembaban Tanah: % kelembaban tanah di sekitar tanaman
- Penyebaran Gejala: Sebaran bercak atau kerusakan daun (umumnya berdasarkan intensitas)

b. Contoh Data

Warna Daun	Bercak Daun	Ukuran Bercak	Tekstur Daun	Kondisi Batang	Kelembaban Tanah (%)	Penyebaran Gejala	Hama	Label Penyakit
Hijau	Tidak Ada	-	Halus	Normal	40	Tidak	Tidak	Tanaman Sehat
Kuning	Ada	Sedang	Kasar	Sedang	30	Sedang	Ada	Hawar Daun Bakteri
Hijau	Ada	Kecil	Halus	Tidak Ada	45	Tidak Ada	Tidak	Bercak Coklat
Coklat	Ada	Besar	Kasar	Tinggi	25	Tinggi	Ada	Penyakit Blast



c. Sumber Penelitian

- 1) Beberapa penelitian yang menggunakan data serupa dalam penerapan Random Forest untuk klasifikasi penyakit tanaman padi antara lain:
- 2) Mutlu & Erenturk (2020): Menggunakan fitur-fitur seperti warna daun, ukuran bercak, dan kondisi batang untuk klasifikasi penyakit daun padi seperti blast dan hawar daun
- 3) Sinha & Arora (2019): Studi tentang penerapan algoritma pembelajaran mesin (ML) seperti Random Forest untuk klasifikasi bercak coklat dan penyakit blast pada daun padi menggunakan gambar daun yang telah diproses menjadi skala abu-abu dan diekstrak fiturnya menggunakan momen intensitas dan tekstur
- 4) Huang et al. (2021): Penelitian menggunakan data reflektansi spektral untuk mendeteksi penyakit padi, termasuk blight, dengan Random Forest sebagai salah satu algoritma utama dalam analisis data lingkungan dan visual

2.7.2 Data Sekunder

Data sekunder adalah data yang diperoleh dari catatan atau laporan pihak lain, bukan dari peneliti langsung, sehingga data ini bisa berasal dari buku, laporan penelitian terdahulu, atau jurnal (Sugiyono., 2013). Data sekunder adalah data yang diperoleh dari sumber yang sudah ada, yang bukan hasil pengumpulan langsung dari lapangan atau eksperimen baru. Data sekunder ini berguna untuk melengkapi atau membandingkan hasil dari data primer, serta menyediakan konteks yang lebih luas dalam memahami pola penyakit tanaman padi.

a. Data Citra (Image Data) Tanaman Padi

Banyak penelitian menggunakan dataset citra daun padi yang diambil dari lapangan atau diunduh dari database publik. Gambar daun yang terinfeksi penyakit seperti blast, hawar daun bakteri, dan bercak coklat diproses untuk ekstraksi fitur seperti tekstur, warna, dan bentuk bercak. Dataset ini sering digunakan untuk pelatihan model Random Forest pada pengenalan penyakit tanaman padi.

Sumber Dataset Citra:

UCI Machine Learning Repository dan Kaggle sering menjadi sumber citra daun padi yang umum digunakan. Jurnal-jurnal seperti Computers and Electronics in Agriculture sering menyediakan deskripsi data yang digunakan dalam eksperimen mereka.

b. Data Ekstraksi Fitur

Peneliti sering mengekstrak fitur dari citra daun, seperti intensitas warna, tekstur (contoh: histogram atau GLCM - Gray-Level Co-occurrence Matrix), dan pola bentuk bercak, yang kemudian dimasukkan ke dalam model machine learning. Data ini mengandung statistik deskriptif dari citra daun yang telah diproses.

Contoh Data Ekstraksi Fitur:

Mean, variance, contrast, entropy dari tekstur daun. Kadar intensitas warna untuk mendeteksi perubahan warna yang disebabkan oleh penyakit. Sumber data sekunder ini biasanya dihasilkan dari penelitian berbasis machine learning, seperti yang ditemukan dalam studi Mutlu & Erenturk (2020) yang menggunakan data GLCM untuk deteksi penyakit.

c. Data Lingkungan (Kondisi Tanah dan Cuaca)

Penyakit pada tanaman padi juga dipengaruhi oleh kondisi lingkungan seperti kelembaban tanah, intensitas cahaya, dan suhu. Dataset ini bisa dikumpulkan dari sensor lingkungan di lapangan atau dari data meteorologi setempat yang diakses oleh peneliti untuk analisis tambahan.

Sumber Data Lingkungan:

Data meteorologi bisa diperoleh dari BMKG (untuk Indonesia) atau layanan cuaca lainnya yang memberikan informasi iklim dan cuaca. Penelitian yang mencakup faktor lingkungan, seperti



dalam studi Huang et al. (2021), menggunakan data ini untuk menilai pengaruh lingkungan pada penyebaran penyakit padi .

d. Dataset Komersial dari Studi Akademik dan Lembaga Riset

Beberapa dataset komersial yang tersedia di International Rice Research Institute (IRRI) atau CABI (Centre for Agriculture and Bioscience International) menyediakan data penyakit tanaman padi yang mencakup informasi tentang jenis penyakit, tingkat keparahan, gejala visual, dan jenis tanah.

Sumber Dataset Komersial:

Data dari IRRI sering digunakan untuk mendukung studi pengendalian penyakit tanaman padi dan dapat diakses melalui kolaborasi atau permintaan izin. Dataset ini umumnya digunakan untuk menguji model Random Forest dalam berbagai studi, dan membantu dalam mengklasifikasikan berbagai penyakit tanaman padi pada tahap eksperimen atau validasi model.

2.8 Data Analysis

Analisis data seharusnya fokus pada pengambilan keputusan bisnis yang relevan, bukan hanya mengumpulkan data dan menghasilkan prediksi (Cassie Kozyrkov.,2020).

3. ANALISA DAN PEMBAHASAN

Pada bagian ini berisi temuan, hasil serta pembahasan dari topik penelitian.

Tabel 2. Hasil Temuan Artikel Relevan

No	Author/ Tahun	Judu Penelitian	Metode Penelitian	Kelebihan Metode yg dipakai	Kekurangan Metode yg dipakai	Faktor Pengaruh Keberhasila n	Kesimpulan
1.	(Annida Purnama wati et al., 2020)	Deteksi Penyakit Daun pada Tanaman Padi Mengguna kan Algoritma Decision Tree, Random Forest, Naïve Bayes, SVM dan KNN	Studi Literatur	Random Forest memiliki akurasi tinggi untuk klasifikasi penyakit daun pada tanaman padi. Algoritma ini mampu menangani dataset besar dan kompleks dengan menghasilkan hasil yang andal dalam klasifikasi.	Model Random Forest rentan terhadap <i>overfitting</i> pada dataset kecil atau terbatas. Hal ini berpotensi menurunkan akurasi pada data uji yang berbeda dari data latih.	Kesesuaian algoritma dengan karakteristi k dataset dan jenis penyakit daun sangat memengaru hi keberhasila n klasifikasi. Pemilihan algoritma yang sesuai seperti KNN atau SVM, yang dapat memberika n hasil lebih optimal dalam kasus tertentu, menjadi penting untuk meningkatk	Penelitian ini menyimpulkan bahwa KNN memiliki akurasi terbaik dalam klasifikasi penyakit daun pada tanaman padi, namun Random Forest tetap menghasilkan performa baik untuk kategori ini.



						an akurasi.	
2.	(Sarifah et al., 2022)	Klasifikasi Penyakit Daun Padi menggunakan Random Forest dan Color Histogram Jurnal Komputasi	Studi Literatur	. Random Forest memiliki stabilitas yang tinggi pada berbagai skenario klasifikasi, yang menjadikannya sangat cocok untuk diagnosis penyakit tanaman. Metode ini mampu mengklasifikasi kan penyakit dengan akurasi yang baik, bahkan ketika terdapat variasi dalam data yang digunakan.	Waktu komputasi yang diperlukan bisa lebih lama, terutama jika dataset yang digunakan sangat besar. Proses pelatihan model juga memerlukan banyak iterasi dan tuning, yang mempengaruhi efisiensi.	Ukuran dan variasi data memainkan peran penting dalam keberhasilan Random Forest. Jika dataset yang digunakan sangat bervariasi atau tidak representatif, maka model tidak akan dapat menghasilkan prediksi yang akurat. Oleh karena itu, kualitas data dan pemilihan fitur yang tepat sangat menentukan keberhasilan metode ini.	Penelitian ini menunjukkan bahwa Random Forest memiliki potensi besar dalam klasifikasi penyakit padi dengan akurasi yang baik, meskipun waktu komputasi yang lama menjadi tantangan yang perlu diatasi untuk aplikasi dalam skala besar.
3.	(Prasetya et al., 2021)	Sistem Pakar untuk Identifikasi Hama Tanaman Padi	Kualitatif Deskriptif	. Random Forest sangat konsisten dalam memberikan akurasi yang tinggi, bahkan dalam kondisi lapangan yang bervariasi. Keuntungannya adalah kemampuan untuk menangani data dengan noise yang tinggi dan tetap memberikan hasil yang andal.	Salah satu kekurangan utama adalah perlu tuning parameter yang optimal. Proses ini dapat memakan waktu dan mengurangi efisiensi aplikasi di lapangan, terutama jika sistemnya tidak terintegrasi dengan baik.	Kondisi lingkungan dan cuaca sangat mempengaruhi efektivitas Random Forest dalam identifikasi hama. Perubahan kondisi cuaca yang cepat atau lingkungan yang tidak stabil bisa mempengaruhi kualitas data yang digunakan untuk	Penelitian ini menunjukkan bahwa Random Forest sangat akurat, aplikasi di lapangan memerlukan penyesuaian parameter yang tepat untuk memastikan sistem berjalan optimal. Selain itu, kondisi lingkungan yang dinamis juga harus dipertimbangkan



						melatih model.	kan.
4.	(Sari et al., 2019)	Deteksi Penyakit Padi Menggunakan Teknologi Android	Studi Literatur	Implementasi yang mudah melalui aplikasi mobile menjadikan Random Forest sangat berguna di lapangan. Petani dapat langsung mengakses dan menggunakan sistem untuk mendeteksi penyakit padi secara cepat.	Sistem ini sangat bergantung pada koneksi internet yang stabil. Tanpa akses internet yang baik, sistem deteksi penyakit bisa terganggu, terutama di daerah yang memiliki konektivitas yang rendah.	Kecepatan aplikasi dan koneksi internet menjadi faktor kunci. Jika aplikasi tidak dapat berfungsi dengan cepat atau terputus karena koneksi yang buruk, maka sistem tidak akan memberikan manfaat maksimal bagi petani.	Penelitian ini menyimpulkan bahwa teknologi berbasis mobile dengan Random Forest sangat membantu petani dalam mendeteksi penyakit padi secara real-time, namun masalah konektivitas menjadi tantangan utama.
5.	(Khairul Rozi et al., 2024)	Klasifikasi Penyakit Tanaman Padi Menggunakan Convolutional Neural Network (CNN) dan Random Forest.	Studi Literatur	Mudah diintegrasikan ke dalam sistem pakar yang sudah ada, menjadikan Random Forest sangat fleksibel dalam berbagai aplikasi pertanian. Hasil diagnosis penyakit yang cepat dan akurat juga menjadi keunggulannya.	Keterbatasan data di lokasi tertentu dapat mengurangi akurasi model. Jika data yang digunakan tidak cukup representatif dari kondisi lapangan, maka diagnosis yang dihasilkan bisa tidak akurat.	Kualitas input data menjadi sangat penting. Tanpa data yang berkualitas dan representatif, model Random Forest tidak akan memberikan hasil yang maksimal.	Penelitian ini menunjukkan bahwa Random Forest memberikan hasil yang cepat dan akurat, namun kualitas data yang masuk sangat mempengaruhi kinerjanya dalam diagnosis penyakit padi.
6.	(Firman Hadi et al., 2023)	Analisis Produktivitas Padi Menggunakan Algoritma Machine Learning Random Forest di Kabupaten Batang Tahun 2018-2022.	Kualitatif Deskriptif	Dapat diakses secara langsung di lapangan melalui perangkat mobile, memungkinkan petani untuk mendeteksi dan mengatasi penyakit padi secara langsung dan efisien.	Pengumpulan data lapangan menjadi faktor yang sangat menentukan. Tanpa data lapangan yang akurat, sistem tidak akan berfungsi optimal, dan kualitas diagnosis akan menurun.	Kapabilitas perangkat mobile yang digunakan dan pengumpulan data lapangan yang akurat sangat mempengaruhi keberhasilan sistem ini.	Sistem berbasis mobile ini terbukti membantu petani mengenali dan mengatasi penyakit padi dengan lebih cepat, namun bergantung pada akurasi pengumpulan data dan



							kualitas perangkat yang digunakan.
7.	(Hamzah et al., 2022)	Identifikasi Penyakit Tanaman Padi dengan Pendekatan AI Menggunakan Random Forest	Studi Eksperimental	. Random Forest efisien dalam pemrosesan data pertanian besar dengan banyak variabel. Algoritma ini dapat mendeteksi berbagai penyakit dengan akurasi yang tinggi dan efisiensi waktu yang baik.	Rentan terhadap kesalahan deteksi jika ada data yang tidak relevan atau noise yang cukup besar. Keterbatasan dalam menangani data yang tidak seimbang.	Kualitas dan konsistensi data sensor yang digunakan serta pemilihan parameter model yang tepat sangat berpengaruh terhadap akurasi hasil deteksi penyakit.	Model ini sangat efektif dalam mendeteksi penyakit tanaman padi, namun penting untuk memiliki data sensor yang berkualitas untuk mendapatkan hasil yang lebih akurat.
8.	(Rudianto et al., 2019)	Analisis Penyakit Padi Menggunakan Data Citra Satelit		. Kemampuan untuk mengolah data besar dari citra satelit dan efektif dalam mengidentifikasi pola geografis yang berkaitan dengan penyebaran penyakit pada tanaman padi. Dapat menganalisis wilayah yang luas dengan cepat.	Kualitas dan resolusi data citra satelit yang terbatas dapat mengurangi akurasi model, terutama dalam deteksi penyakit di wilayah yang tidak terjangkau. Diperlukan perangkat keras yang mumpuni untuk pemrosesan data besar.	Kualitas dan resolusi data citra satelit yang digunakan sangat berpengaruh pada akurasi model. Selain itu, teknik preprocessing citra juga menjadi faktor penting untuk meningkatkan kualitas data yang akan diproses.	Penelitian ini menyimpulkan bahwa penggunaan Random Forest dalam pemetaan penyakit padi berbasis citra satelit memiliki potensi besar, namun tantangan utama adalah kualitas citra yang terbatas dan ketersediaan data yang relevan.
9.	(Zulkifli et al., 2020)	Sistem Pendeteksi Penyakit Tanaman Padi dengan Algoritma Random Forest		Memberikan diagnosa yang cepat dengan akurasi tinggi. Dapat mengidentifikasi penyakit padi dengan ketepatan tinggi berdasarkan data yang diberikan.	Memerlukan dataset yang besar dan berkualitas. Performa model dapat berkurang jika data tidak cukup representatif.	Ketersediaan dataset yang lengkap dan relevansi fitur yang dipilih sangat berpengaruh terhadap akurasi model.	Random Forest terbukti efektif dalam mendeteksi penyakit padi pada skala yang lebih luas, namun pemilihan dan pengolahan data yang tepat sangat memengaruhi



							hasilnya.
10.	(Dirgahayu et al., 2015)	Model Pertumbuhan Padi Menggunakan Data EVI MODIS	Kualitatif Deskriptif	Random Forest sangat efektif dalam mengklasifikasi fase pertumbuhan padi, terutama ketika menggunakan data citra satelit seperti EVI. Metode ini tidak hanya mampu menangani dataset besar tetapi juga memberikan hasil yang stabil dan konsisten dalam memprediksi fase pertumbuhan tanaman.	Salah satu tantangan terbesar adalah sensitivitas terhadap data noise. Keberhasilan klasifikasi sangat bergantung pada kebersihan dan kualitas data. Apabila data yang digunakan mengandung noise atau kesalahan, maka hasil klasifikasi akan terganggu.	Konsistensi data sangat penting untuk keberhasilan Random Forest. Jika data EVI yang digunakan tidak konsisten atau mengandung banyak noise, hasil klasifikasi bisa menjadi tidak akurat. Selain itu, frekuensi dan akurasi pemantauan data sangat mempengaruhi hasil analisis.	Model ini mengungkapkan bahwa meskipun Random Forest efektif dalam mengklasifikasi fase pertumbuhan padi, kualitas dan konsistensi data sangat penting untuk memperoleh hasil yang akurat dan berguna bagi pengelolaan pertanian.

4. KESIMPULAN

Metode Random Forest terbukti efektif untuk klasifikasi penyakit pada tanaman padi, menunjukkan akurasi tinggi dan stabilitas dalam berbagai skenario. Keunggulan utama Random Forest adalah kemampuannya dalam menangani dataset yang kompleks dan beragam, membuatnya cocok untuk diagnosa penyakit pada tanaman padi. Namun, penelitian juga menunjukkan beberapa keterbatasan, seperti ketergantungan pada kualitas dan representasi data yang digunakan, serta potensi waktu komputasi yang lebih lama ketika mengolah dataset besar. Dalam perbandingan dengan algoritma lain seperti SVM dan Decision Tree, Random Forest umumnya unggul dalam akurasi, tetapi memerlukan tuning parameter yang cermat untuk performa optimal. Implementasi pada aplikasi lapangan juga menghadapi tantangan teknis, terutama terkait kualitas konektivitas dan ketersediaan perangkat mobile bagi petani. Secara keseluruhan, Random Forest memiliki potensi besar untuk mendukung teknologi pertanian cerdas yang dapat membantu meningkatkan produktivitas pertanian padi di Indonesia.

REFERENCES

- Kusuma, D. A., & Satria, H. (2023). Implementasi Algoritma Random Forest Regression untuk Memprediksi Hasil Panen Padi di Desa Minanga. ResearchGate. https://www.researchgate.net/publication/375885030_Implementasi_Algoritma_Random_Forest_Regression_untuk_Memprediksi_Hasil_Panen_Padi_di_Desa_Minanga
- Fajar, I., & Harsono, B. (2023). Penerapan Algoritma Klasifikasi untuk Sistem Pakar. Jurnal Infotekjar, 10(3), 55-67. <https://www.jurnal.uisu.ac.id/index.php/infotekjar/article/view/2934>
- Wahono, R. S. (2016). Systematic Literature Review: Pengantar, Tahapan, dan Studi Kasus. Romisatriawahono.net. <https://romisatriawahono.net/2016/05/15/systematic-literature-review-pengantar-tahapan-dan-studi-kasus/>



- Putra, H. A., & Widi, T. (2023). Prediksi Kelayakan Kredit Menggunakan Machine Learning. *Jurnal Komputasi*, 5(2), 45-56. <https://komputasi.fmipa.unila.ac.id/index.php/komputasi/article/download/216/216/250>
- Sari, E. T., & Agustina, R. (2023). Klasifikasi Citra Digital Menggunakan Metode Random Forest. *Repository IPB*. <https://repository.ipb.ac.id/handle/123456789/105331>
- Susanto, I., & Lestari, A. (2022). Deteksi Penyakit Tanaman Menggunakan Metode Random Forest. *Jurnal Mnemonic*, 4(1), 78-85. <https://ejournal.itn.ac.id/index.php/mnemonic/article/download/10187/6119>
- Putri, L. D., & Santoso, B. (2021). Penggunaan Machine Learning dalam Prediksi Penjualan. *Jurnal Variansi*, 2(1), 90-102. <https://jurnalvariansi.unm.ac.id/index.php/variansi/article/download/31/11>
- Fahmi, D., & Ahmad, Y. (2022). Aplikasi Pendeteksian Objek Menggunakan Algoritma Klasifikasi. *Xjournals*. <https://xjournals.com/collections/articles/Article?qt=pNGIy7BUHXUapgsdP1tZP+BzEuALcQXlortgKGdPcDk=>
- Suryani, R., & Wijaya, T. (2023). Penggunaan Random Forest dalam Sistem Informasi Geografis. *Jurnal Komputasi*, 7(2), 123-130. <https://komputasi.fmipa.unila.ac.id/index.php/komputasi/article/view/216>
- Haryanto, B., & Sudrajat, D. (2022). Analisis Data Kesehatan Menggunakan Metode Klasifikasi. *Jurnal Mnemonic*, 5(3), 57-64. <https://ejournal.itn.ac.id/index.php/mnemonic/article/view/10187>
- Gunawan, A., & Pratama, K. (2023). Metode Random Forest untuk Analisis Data Cuaca. *Jurnal Elipsoida*, 3(2), 76-83. <https://ejournal2.undip.ac.id/index.php/elipsoida/article/view/19023>
- Iskandar, Z., & Ramadhan, S. (2023). Rice Leaf Diseases Detection Using Machine Learning. *ResearchGate*. https://www.researchgate.net/publication/377215046_Rice_Leaf_Diseases_Detection_Using_Machine_Learning
- Rosyani, P. (2024). Implementasi Random Forest dalam Prediksi Risiko Kesehatan. *Google Scholar*. https://scholar.google.com/scholar?hl=id&as_sdt=0%2C5&q=perani+rosyani+Random+foorest&btnG=#d=gs_qabs&t=1731391934347&u=%23p%3DsusWdXPcHpgJ
- Rosyani, P. (2024). Penerapan Algoritma Klasifikasi untuk Diagnosa Penyakit. *Google Scholar*. https://scholar.google.com/scholar?hl=id&as_sdt=0%2C5&q=perani+rosyani+Random+foorest&btnG=#d=gs_qabs&t=1731392007279&u=%23p%3D9-yOVrsnpMEJ
- Ahmad, F. (2023). Prediksi Kualitas Air Menggunakan Random Forest. *Academia*. <https://www.academia.edu/download/89847530/pdf.pdf>
- Suryono, B. A. (2021). Sistem Pakar Untuk Analisis Penyakit Kulit. *Jurnal Sisfotek*, 3(2), 45-52. <https://www.jsisfotek.org/index.php/JSisfotek/article/view/144>
- Hadi, R. P., & Rahayu, M. (2022). Implementasi Machine Learning untuk Pengelompokan Data Pasien. *Proceeding STAIN*, 4(1), 98-105. <https://proceeding.unpkediri.ac.id/index.php/stains/article/view/2884>
- Sutanto, D., & Ratna, A. (2021). Analisis Data Lahan Pertanian Menggunakan Metode Random Forest. *Jurnal Agv*, 6(2), 115-124. <https://ppjp.ulm.ac.id/journals/index.php/agv/article/view/1246>
- Kurniawan, T., & Diah, M. (2023). Sistem Prediksi Pertanian Menggunakan Algoritma Random Forest. *Jurnal Syntax*, 5(3), 67-74. <https://journal.unsika.ac.id/index.php/syntax/article/view/9945>
- Wardani, E. S., & Nanda, R. (2024). Penerapan Machine Learning untuk Deteksi Penyakit Tanaman. *Jurnal RTP*, 7(1), 44-51. <https://jurnal.usk.ac.id/RTP/article/view/31147>
- Wijaya, R. S., & Tania, P. (2024). Penggunaan Random Forest dalam Sistem Informasi Geospasial. *Jurnal Geoid*, 2(2), 102-108. <https://iptek.its.ac.id/index.php/geoid/article/view/828>