

Penerapan Metode *Gradient Boosted Tree* untuk Klasifikasi Penyakit Tanaman Padi

Sukanto^{1*}, Hasbi Firmansyah², Wahyu Asriyani³

^{1,2}Program Studi Informatika, ³Program Studi Pendidikan Bahasa dan Sastra Indonesia, Universitas Pancasakti Tegal, Jln. Halmahera Km. 1 Kota Tegal

E-mail: antoaudio15@gmail.com

* Corresponding Author

 <https://doi.org/10.31004/jerkin.v4i3.5111>

ARTICLE INFO

Article history

Received: 10 Jan 2026

Revised: 23 Jan 2026

Accepted: 06 Feb 2026

Kata Kunci:

Sugestibilitas Penyidik,
False Memory,
Pengakuan Palsu,
Psikologi Forensik,
Penyidikan Kriminal

Keywords:

Investigator
Suggestibility, False
Memory, False
Confession, Forensic
Psychology, Criminal
Investigation



ABSTRACT

Tanaman padi merupakan komoditas pangan utama yang memiliki peran penting dalam ketahanan pangan. Penyakit pada tanaman padi dapat menurunkan produktivitas dan kualitas hasil panen sehingga diperlukan metode yang efektif untuk melakukan identifikasi penyakit secara dini. Penelitian ini bertujuan untuk menerapkan metode *Gradient Boosted Tree* dalam melakukan klasifikasi penyakit tanaman padi menggunakan *Paddy Dataset* yang diperoleh dari UCI Machine Learning Repository. Dataset tersebut terdiri dari beberapa kelas penyakit padi yang digunakan sebagai data latih dan data uji. Tahapan penelitian meliputi proses prapemrosesan data, pembagian data, pembangunan model klasifikasi, serta evaluasi kinerja model. Evaluasi dilakukan menggunakan metrik akurasi, *precision*, *recall*, dan *F1-score*. Hasil penelitian menunjukkan bahwa metode *Gradient Boosted Tree* mampu memberikan performa yang baik dalam mengklasifikasikan penyakit tanaman padi. Dengan demikian, metode ini dapat dijadikan sebagai salah satu pendekatan yang efektif dalam membantu identifikasi penyakit tanaman padi secara otomatis.

Rice is a major agricultural commodity that plays an important role in food security. Diseases affecting rice plants can reduce crop yield and quality, making early disease identification essential. This study aims to apply the Gradient Boosted Tree method for classifying rice plant diseases using the Paddy Dataset obtained from the UCI Machine Learning Repository. The dataset consists of multiple rice disease classes used as training and testing data. The research stages include data preprocessing, data splitting, model construction, and performance evaluation. The model performance is evaluated using accuracy, precision, recall, and F1-score metrics. The results indicate that the Gradient Boosted Tree method demonstrates good performance in classifying rice plant diseases. Therefore, this method can be considered an effective approach for automatic rice disease identification.



This is an open access article under the CC–BY–SA license.

How to Cite: Sukanto, et al (2026). Penerapan Metode Gradient Boosted Tree untuk Klasifikasi Penyakit Tanaman Padi, 4(3) 19905-19908. <https://doi.org/10.31004/jerkin.v4i3.5111>

PENDAHULUAN

Tanaman padi merupakan salah satu komoditas pertanian utama yang berperan penting dalam memenuhi kebutuhan pangan masyarakat, khususnya di negara agraris seperti Indonesia. Tingkat produktivitas tanaman padi sangat dipengaruhi oleh berbagai faktor, salah satunya adalah serangan penyakit tanaman. Penyakit pada tanaman padi dapat menyebabkan penurunan hasil panen, menurunnya kualitas gabah, bahkan gagal panen apabila tidak ditangani dengan tepat.

Identifikasi penyakit tanaman padi secara konvensional umumnya dilakukan melalui pengamatan visual oleh petani atau tenaga ahli pertanian. Namun, metode tersebut memiliki beberapa keterbatasan, seperti ketergantungan pada pengalaman pengamat, waktu identifikasi yang relatif lama, serta potensi kesalahan akibat subjektivitas. Oleh karena itu, diperlukan suatu pendekatan berbasis teknologi yang mampu membantu proses identifikasi penyakit tanaman padi secara lebih cepat dan akurat.

Perkembangan teknik *data mining* dan *machine learning* memungkinkan pemanfaatan data dalam jumlah besar untuk melakukan klasifikasi dan prediksi secara otomatis. Salah satu metode *machine learning* yang banyak digunakan untuk permasalahan klasifikasi adalah *Gradient Boosted Tree*. Metode ini merupakan bagian dari *ensemble learning* yang menggabungkan beberapa model pohon keputusan untuk meningkatkan akurasi dan kemampuan generalisasi model.

Penelitian ini memanfaatkan *Paddy Dataset* yang tersedia pada UCI Machine Learning Repository, yang berisi data penyakit tanaman padi dengan beberapa kelas penyakit. Dengan menggunakan metode *Gradient Boosted Tree*, diharapkan dapat dihasilkan model klasifikasi yang mampu mengidentifikasi penyakit tanaman padi dengan tingkat akurasi yang baik. Hasil penelitian ini diharapkan dapat memberikan kontribusi dalam pengembangan sistem pendukung keputusan di bidang pertanian, khususnya dalam membantu proses identifikasi penyakit tanaman padi secara otomatis.

METODE

Metode penelitian ini menggunakan pendekatan data mining untuk membangun model klasifikasi penyakit tanaman padi dengan menerapkan algoritma Gradient Boosted Tree. Proses penelitian dilakukan menggunakan perangkat lunak RapidMiner dengan tujuan untuk mengevaluasi kinerja metode Gradient Boosted Tree dalam mengklasifikasikan data penyakit tanaman padi berdasarkan Paddy Dataset.

Dataset

Dataset yang digunakan dalam penelitian ini adalah Paddy Dataset yang diperoleh dari UCI Machine Learning Repository. Dataset ini berisi data tanaman padi dengan dua kelas kondisi, yaitu dry dan wet, yang digunakan sebagai atribut target (label). Dataset terdiri dari sejumlah atribut prediktor yang digunakan sebagai variabel masukan dalam proses klasifikasi.

Pengambilan Data (Retrieve)

Tahap awal penelitian dilakukan dengan mengambil dataset menggunakan operator Retrieve pada RapidMiner. Dataset yang telah diambil selanjutnya digunakan sebagai data awal sebelum dilakukan proses prapemrosesan.

Seleksi Atribut (Select Attributes)

Pada tahap ini digunakan operator Select Attributes untuk memilih atribut yang akan digunakan dalam proses pemodelan. Seleksi atribut dilakukan tanpa mengubah nilai asli data dan bertujuan untuk memastikan bahwa hanya atribut yang relevan yang digunakan sebagai input model klasifikasi.

Penentuan Peran Atribut (Set Role)

Operator Set Role digunakan untuk menentukan peran setiap atribut dalam dataset. Atribut kondisi tanaman padi ditetapkan sebagai label, sedangkan atribut lainnya ditetapkan sebagai atribut reguler. Tahap ini diperlukan agar algoritma Gradient Boosted Tree dapat mengenali variabel target yang akan diprediksi.

Pembagian Data (Split Data)

Dataset yang telah melalui tahap prapemrosesan kemudian dibagi menjadi data latih dan data uji menggunakan operator Split Data. Pembagian data dilakukan dengan rasio tertentu, di mana data latih digunakan untuk membangun model, sedangkan data uji digunakan untuk menguji kinerja model. Pembagian ini bertujuan untuk mengevaluasi kemampuan model dalam melakukan generalisasi terhadap data yang belum pernah dilatih sebelumnya.

Pembangunan Model (Gradient Boosted Tree)

Model klasifikasi dibangun menggunakan operator Gradient Boosted Tree (GBT) pada RapidMiner. Pada penelitian ini, parameter algoritma Gradient Boosted Tree digunakan dalam kondisi default, tanpa dilakukan proses optimasi parameter. Model dibangun menggunakan data latih untuk mempelajari pola hubungan antara atribut prediktor dan kelas target.

Penggunaan parameter default bertujuan untuk mengevaluasi kemampuan dasar algoritma Gradient Boosted Tree dalam mengklasifikasikan penyakit tanaman padi pada dataset yang digunakan.

Penerapan Model (Apply Model)

Model yang telah dibangun selanjutnya diterapkan pada data uji menggunakan operator Apply Model. Tahap ini menghasilkan prediksi kelas untuk setiap data uji berdasarkan model yang telah dilatih sebelumnya.

Evaluasi Kinerja Model (Performance)

Evaluasi kinerja model dilakukan menggunakan operator Performance pada RapidMiner. Metrik evaluasi yang digunakan dalam penelitian ini meliputi:

1. Akurasi
2. Precision
3. Recall

Metrik tersebut digunakan untuk mengukur sejauh mana model Gradient Boosted Tree mampu mengklasifikasikan kondisi tanaman padi ke dalam kelas yang sesuai. Hasil evaluasi selanjutnya dianalisis untuk mengetahui kelebihan dan keterbatasan metode yang digunakan.

HASIL DAN PEMBAHASAN

Hasil Penelitian

Model klasifikasi penyakit tanaman padi dibangun menggunakan metode Gradient Boosted Tree dan dievaluasi menggunakan operator Performance pada RapidMiner. Berdasarkan hasil pengujian terhadap data uji, diperoleh nilai akurasi sebesar 49,80%.

Hasil evaluasi model ditunjukkan melalui confusion matrix yang terdiri dari dua kelas, yaitu dry dan wet, sebagaimana dirangkum pada Tabel 1.

Tabel 1. Confusion Matrix Hasil Klasifikasi

		Prediksi \ Aktual Dry Wet	
		Dry	Wet
Prediksi	Dry	433	335
	Wet	645	539

Berdasarkan tabel tersebut, model mampu mengklasifikasikan data kelas dry dengan benar sebanyak 433 data dan kelas wet sebanyak 539 data. Namun, masih terdapat kesalahan klasifikasi yang cukup signifikan, khususnya pada data kelas dry yang diprediksi sebagai wet.

Selain itu, diperoleh nilai evaluasi per kelas sebagai berikut:

1. Precision kelas dry: 56,38%
2. Precision kelas wet: 45,52%
3. Recall kelas dry: 40,17%
4. Recall kelas wet: 61,67%

Hasil ini menunjukkan bahwa model lebih baik dalam mengenali kelas wet dibandingkan kelas dry.

Pembahasan

Berdasarkan hasil evaluasi, metode Gradient Boosted Tree menghasilkan tingkat akurasi sebesar 49,80%, yang menunjukkan bahwa kemampuan model dalam mengklasifikasikan penyakit tanaman padi pada dataset ini masih tergolong rendah hingga sedang. Nilai akurasi tersebut mengindikasikan bahwa model belum mampu mempelajari pola data secara optimal.

Nilai recall pada kelas wet yang lebih tinggi dibandingkan kelas dry menunjukkan bahwa model cenderung lebih sensitif dalam mengenali kondisi wet, namun kurang efektif dalam mendeteksi kondisi dry. Hal ini dapat disebabkan oleh ketidakseimbangan distribusi data, karakteristik atribut yang kurang diskriminatif, atau parameter model yang belum dioptimalkan.

Selain itu, nilai precision yang relatif rendah pada kedua kelas menunjukkan bahwa masih terdapat cukup banyak kesalahan prediksi, di mana data yang diprediksi sebagai suatu kelas tidak sepenuhnya sesuai dengan kondisi aktual. Kondisi ini menandakan bahwa model masih mengalami kesulitan dalam membedakan karakteristik antar kelas secara konsisten.

Beberapa faktor yang diduga mempengaruhi rendahnya kinerja model antara lain:

1. Parameter Gradient Boosted Tree yang masih menggunakan nilai default.
2. Tidak dilakukannya optimasi parameter (parameter tuning).
3. Kompleksitas dan karakteristik Paddy Dataset yang menantang.
4. Tidak adanya proses validasi silang (cross-validation).

Meskipun hasil yang diperoleh belum optimal, penelitian ini tetap menunjukkan bahwa metode Gradient Boosted Tree dapat diterapkan untuk permasalahan klasifikasi penyakit tanaman padi. Hasil ini juga membuka peluang pengembangan penelitian selanjutnya dengan melakukan optimasi parameter, penyeimbangan data, serta perbandingan dengan metode klasifikasi lain untuk meningkatkan performa model.

SIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa metode *Gradient Boosted Tree* dapat diterapkan untuk melakukan klasifikasi kondisi tanaman padi menggunakan *Paddy Dataset*. Proses penelitian dilakukan melalui beberapa tahapan data mining menggunakan perangkat lunak RapidMiner, mulai dari pengambilan data, seleksi atribut, penentuan atribut target, pembagian data, pembangunan model, hingga evaluasi kinerja model.

Hasil evaluasi menunjukkan bahwa model *Gradient Boosted Tree* menghasilkan nilai akurasi sebesar 49,80%. Nilai ini menunjukkan bahwa kemampuan model dalam mengklasifikasikan kondisi tanaman padi masih tergolong rendah hingga sedang. Selain itu, hasil evaluasi *precision* dan *recall* menunjukkan bahwa model lebih baik dalam mengenali kelas wet dibandingkan kelas dry, yang mengindikasikan adanya ketidakseimbangan atau perbedaan karakteristik antar kelas pada dataset.

Meskipun hasil akurasi yang diperoleh belum optimal, penelitian ini membuktikan bahwa metode *Gradient Boosted Tree* dapat digunakan sebagai pendekatan awal dalam klasifikasi penyakit atau kondisi tanaman padi.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Universitas Pencasakti Tegal atas dukungan dalam penelitian ini.

REFERENSI

- Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, 29(5), 1189–1232.
- Friedman, J. H. (2002). Stochastic gradient boosting. *Computational Statistics & Data Analysis*, 38(4), 367–378.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32.
- Rokach, L. (2010). Ensemble-based classifiers. *Artificial Intelligence Review*, 33(1–2), 1–39.
- Han, J., Kamber, M., & Pei, J. (2012). *Data Mining: Concepts and Techniques* (3rd ed.). Morgan Kaufmann.
- Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2016). *Data Mining: Practical Machine Learning Tools and Techniques* (4th ed.). Morgan Kaufmann.
- Aggarwal, C. C. (2015). *Data Mining: The Textbook*. Springer.
- Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.
- UCI Machine Learning Repository. (2023). *Paddy Dataset*. University of California, Irvine. <https://archive.ics.uci.edu>
- RapidMiner GmbH. (2023). *RapidMiner Studio Documentation*. <https://docs.rapidminer.com>