

Implementasi Algoritma *Machine Learning* untuk Deteksi Penyakit Daun Tebu: Analisis Perbandingan Kinerja

Agus Wantoro^{1*}, Dwi Feriyanto², Arry Verdian³, Rakhmat Dedi Gunawan⁴

¹Teknik Informatika, Fakultas Teknologi dan Informatika, Universitas Aisyah Pringsewu

²Rekayasa Perangkat Lunak, Fakultas Teknologi dan Informatika, Universitas Aisyah Pringsewu

³Pendidikan Informatika, Fakultas Pendidikan, STKIP Rosalia Metro

⁴Fakultas Teknik dan Ilmu Komputer, Universitas Teknokrat Indonesia

Email: dwiferiyanto@aisyahuniversity.ac.id

Abstrak

Tebu merupakan komoditas perkebunan penting yang menjadi bahan baku utama industri gula nasional. Produktivitas tebu sangat dipengaruhi oleh serangan penyakit daun seperti karat daun, bercak daun, dan mosaik tebu, yang dapat menurunkan hasil panen secara signifikan. Identifikasi penyakit secara manual oleh tenaga ahli masih menjadi kendala karena sifatnya subjektif, memerlukan waktu lama, serta sulit dilakukan pada area perkebunan yang luas. Untuk mengatasi hal tersebut, diperlukan teknologi seperti algoritma Machine Learning (ML). Tujuan penelitian ini bertujuan membandingkan kinerja dan kompleksitas komputasi beberapa algoritma klasifikasi citra digital dalam mendeteksi penyakit daun tebu secara otomatis. Metode penelitian meliputi pengumpulan dataset citra daun tebu, pra-pemrosesan, ekstraksi citra, serta penerapan beberapa algoritma klasifikasi seperti Naive Bayes (NB), Support Vector Machine (SVM), k-Nearest Neighbor (KNN), Random Forest (RF), dan Neural Network (NN). Setiap algoritma dievaluasi berdasarkan akurasi, precision, recall, serta kompleksitas komputasi yang diukur melalui waktu komputasi (train). Hasil pengujian kinerja menunjukkan bahwa algoritma NN memberikan akurasi tertinggi, sedangkan algoritma Random Forest (RF) lebih efisien secara komputasi. Temuan ini akan memberikan kontribusi penting dalam pengembangan sistem deteksi penyakit daun tebu berbasis citra digital yang akurat dan efisien, sebagai bagian dari penerapan pertanian presisi (precision agriculture) di Indonesia.

Kata kunci: Tebu, Penyakit daun, Klasifikasi citra, *Machine learning*

Abstract

Sugarcane is an important plantation commodity that serves as the main raw material for the national sugar industry. Sugarcane productivity is significantly affected by leaf diseases such as rust, leaf spot, and sugarcane mosaic, which can significantly reduce yields. Manual disease identification by experts remains a challenge due to its subjective nature, time-consuming, and difficult to implement on large plantation areas. To overcome this, technologies such as Machine Learning (ML) algorithms are needed. The purpose of this study is to compare the performance and computational complexity of several digital image classification algorithms in automatically detecting sugarcane leaf diseases. The research methods include collecting sugarcane leaf image datasets, pre-processing, image extraction, and the application of several classification algorithms such as Naive Bayes (NB), Support Vector Machine (SVM), k-Nearest Neighbor (KNN), Random Forest (RF), and Neural Network (NN). Each algorithm was evaluated based on accuracy, precision, recall, and computational complexity as measured by computational time (training). The performance test results showed that the NN algorithm provided the highest accuracy, while the Random Forest (RF) algorithm was more computationally efficient. This finding will provide an important contribution in the development of an accurate and efficient digital image-based sugarcane leaf disease detection system, as part of the implementation of precision agriculture in Indonesia.

Keywords: Sugarcane, Leaf disease, Image classification, Machine learning

PENDAHULUAN

Tebu (*Saccharum officinarum* L.) merupakan salah satu komoditas perkebunan strategis yang menjadi bahan baku utama dalam industri gula nasional (Kumar *et al.*, 2022). Produktivitas tebu sangat dipengaruhi oleh kondisi kesehatan tanaman, khususnya pada bagian daun yang berperan penting dalam proses fotosintesis (Huang *et al.*, 2020). Serangan penyakit daun seperti Penyakit Karat Daun (*Puccinia melanocephala*), Penyakit Bercak Daun (*Cercospora longipes*), dan Penyakit Mozaik Tebu (Sugarcane Mosaic Virus) sering menimbulkan penurunan kualitas dan kuantitas hasil panen. Gejala-gejala penyakit ini umumnya ditandai oleh perubahan warna, pola bercak, dan bentuk daun yang khas, sehingga dapat diidentifikasi melalui citra digital (Chen *et al.*, 2025).

Dalam praktik di lapangan, identifikasi penyakit daun tebu masih banyak dilakukan secara manual oleh tenaga ahli atau petani, yang sangat bergantung pada pengalaman dan kondisi pencahayaan saat pengamatan (Hidayah, 2020). Pendekatan ini bersifat subjektif, memakan waktu, dan tidak efisien, terutama untuk area perkebunan yang luas. Oleh karena itu, diperlukan metode otomatis berbasis pengolahan citra digital dan Machine Learning yang mampu mengidentifikasi penyakit secara cepat, akurat, dan konsisten.

Namun, hingga saat ini belum terdapat standar algoritma klasifikasi yang optimal untuk mendeteksi penyakit daun tebu. Berbagai penelitian telah menerapkan algoritma seperti Naive Bayes (NB), Support Vector Machine (SVM), k-Nearest Neighbors (KNN), Random Forest (RF), dan Neural Network (CNN), tetapi hasil akurasi dan efisiensi komputasi yang diperoleh masih bervariasi tergantung pada karakteristik dataset, teknik ekstraksi fitur, dan parameter model yang digunakan.

Penerapan teknologi klasifikasi citra pada bidang pertanian, khususnya tanaman tebu, memiliki urgensi yang tinggi untuk mendukung konsep pertanian presisi (precision agriculture). Dengan sistem klasifikasi yang akurat dan efisien, deteksi dini penyakit daun dapat dilakukan menggunakan citra digital. Hal ini tidak hanya menghemat biaya dan waktu, tetapi juga memungkinkan pengambilan keputusan cepat untuk pengendalian penyakit, sehingga dapat meningkatkan produktivitas dan keberlanjutan industri gula nasional. Penelitian terdahulu menunjukkan bahwa algoritma berbasis deep learning seperti CNN mampu memberikan akurasi tinggi dalam klasifikasi penyakit daun pada berbagai tanaman, seperti padi, tomat, dan jagung (Srivastava *et al.*, 2020). Namun, algoritma ini sering membutuhkan sumber daya komputasi besar dan waktu pelatihan yang lama.

Sebaliknya, algoritma ML konvensional seperti SVM atau Random Forest menawarkan kecepatan yang lebih baik dengan akurasi yang kompetitif, terutama bila didukung teknik ekstraksi fitur warna, tekstur, dan bentuk yang tepat (Thite *et al.*, 2024). Sebagian besar studi terdahulu berfokus hanya pada perbandingan akurasi, tanpa mempertimbangkan kompleksitas komputasi dan efisiensi model, yang menjadi faktor penting untuk implementasi di perangkat lapangan berdaya rendah

Berdasarkan kajian literatur, masih terdapat kesenjangan penelitian (research gap) dalam beberapa hal yaitu (a) belum adanya kajian komprehensif yang membandingkan performa akurasi dan kompleksitas komputasi dari berbagai algoritma klasifikasi pada citra daun tebu secara bersamaan. (b) minimnya penelitian yang mengukur efisiensi model dari sisi waktu pelatihan secara kuantitatif. (c) kurangnya analisis mengenai trade-off antara akurasi dan efisiensi komputasi untuk menentukan model terbaik yang seimbang antara kinerja dan implementabilitas di lapangan

Penelitian ini bertujuan untuk (a) Perbandingan algoritma klasifikasi penyakit daun tebu berbasis citra digital menggunakan algoritma (NB, k-NN, SVM, NN, dan RF). (b) Membandingkan akurasi, dan waktu komputasi dari masing-masing algoritma. (c) Menentukan algoritma paling optimal yang mampu memberikan keseimbangan antara akurasi tinggi dan efisiensi komputasi

METODE

Penelitian ini dirancang dengan pendekatan eksperimen kuantitatif menggunakan citra digital daun tebu yang diklasifikasikan berdasarkan jenis penyakitnya. Tahapan penelitian secara umum meliputi pengumpulan data, pra-pemrosesan citra, ekstraksi fitur, penerapan algoritma klasifikasi, evaluasi kinerja, dan analisis hasil. Alur penelitian digambarkan pada Gambar 1.



Gambar 1. Tahapan penelitian

Pengumpulan Data Citra

Tahap awal meliputi pengumpulan citra daun tebu yang mewakili berbagai kondisi seperti citra daun sehat dan daun yang terinfeksi penyakit. Pengumpulan data citra diambil dari repositori <https://www.kaggle.com>. Kami menggunakan 2521 yang terbagi kedalam lima penyakit yaitu healthy, Mosaic, RedRot, Rust, dan Yellow (Li *et al.*, 2024). Setiap citra kemudian diberi label sesuai jenis penyakit berdasarkan identifikasi pakar atau deskripsi referensi ilmiah seperti Gambar 2.



Gambar 2. Jenis penyakit daun tebu

Pra-pemrosesan Citra (Preprocessing)

Tahap ini bertujuan untuk meningkatkan kualitas citra agar lebih optimal dalam proses analisis. Beberapa langkah yang dilakukan antara lain:

1. Resizing yaitu menyeragamkan ukuran citra agar sesuai dengan input model.
2. Filtering dan Noise Removal untuk menghilangkan noise
3. Segmentasi Daun yaitu memisahkan objek daun dari latar belakang menggunakan thresholding
4. Augmentasi Data yaitu melakukan rotasi, flipping, dan perubahan kecerahan untuk memperbanyak variasi data dan menghindari overfitting.

Ekstraksi Fitur

Ekstraksi fitur dilakukan untuk memperoleh representasi numerik dari citra yang akan digunakan oleh algoritma klasifikasi.

Penerapan Algoritma Klasifikasi

Beberapa algoritma akan diterapkan dan dibandingkan untuk mengidentifikasi penyakit daun tebu, antara lain:

1. Support Vector Machine (SVM)
2. Naive Bayes (NB)
3. k-Nearest Neighbor (k-NN)
4. Random Forest (RF)

5. Neural Network (NN)

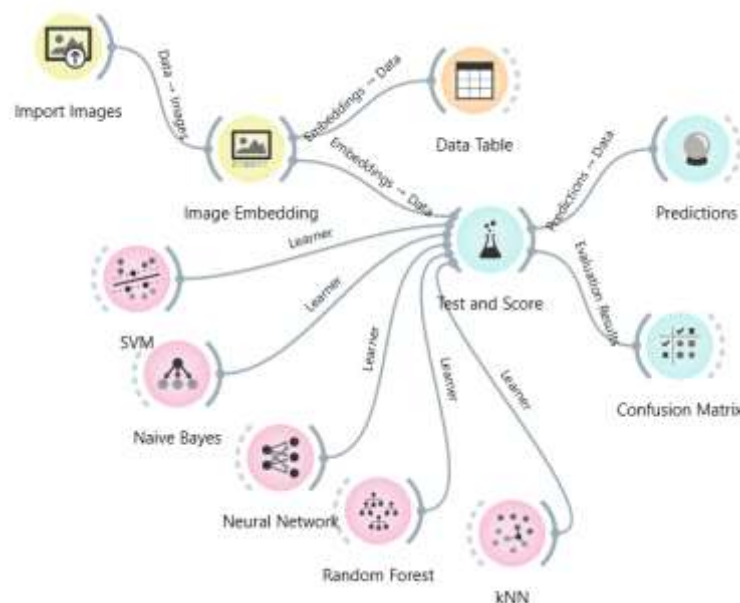
Setiap algoritma akan dilatih dan diuji menggunakan data yang telah diproses, dengan pembagian data pada data latih (training) dan uji (testing).

Evaluasi Kinerja

Evaluasi dilakukan dengan dua perspektif utama yaitu kinerja klasifikasi berupa nilai *Accuracy*, *Precision*, dan *Recall*. Selain itu, kami melakukan analisis perbandingan kompleksitas komputasi yaitu mengevaluasi kinerja berdasarkan waktu pelatihan (training time) dan waktu prediksi (*inference time*). Pendekatan ini bertujuan untuk menemukan model yang tidak hanya akurat, tetapi juga efisien secara komputasi, sehingga dapat diimplementasikan pada perangkat dengan sumber daya terbatas

HASIL DAN PEMBAHASAN

Kami menggunakan tools Orang untuk melakukan perbandingan algoritma ML. Tools ini mampu mengintegrasikan berbagai algoritma machine learning dan data mining (seperti klasifikasi, klusterisasi, dan regresi). Orange memungkinkan pengguna untuk membangun dan mengevaluasi model prediktif dengan mudah (Akkaya & Turgay, 2024). Gambar 3 menampilkan desain pengujian algoritma.



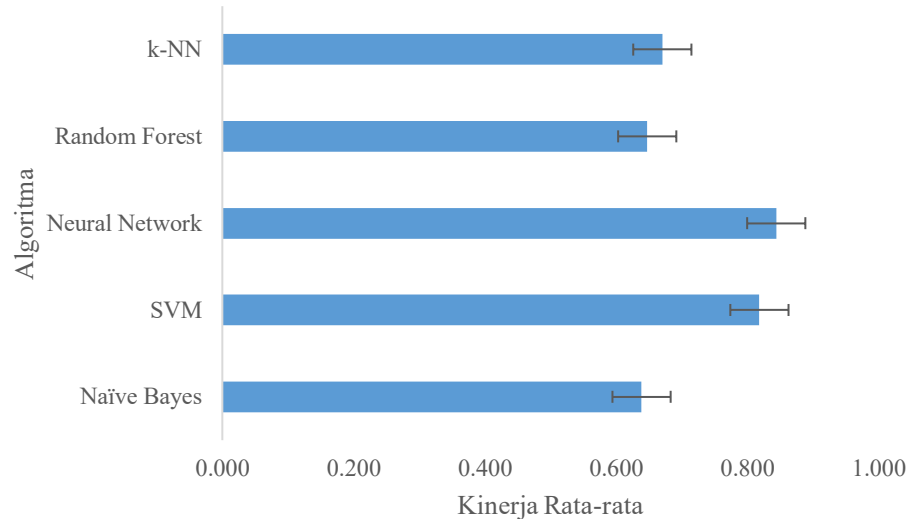
Gambar 3. Desain model pengujian algoritma ML

Kami menggunakan cross-validation (k-fold=5) dengan komposisi data training 80% dan data testing 20%. Hasil temuan kami, kompisis ini memiliki kinerja optimal. Hasil perbandingan kinerja ditampilkan dalam Tabel 1.

Tabel 1. Perbandingan Akurasi, Precision, dan Recall

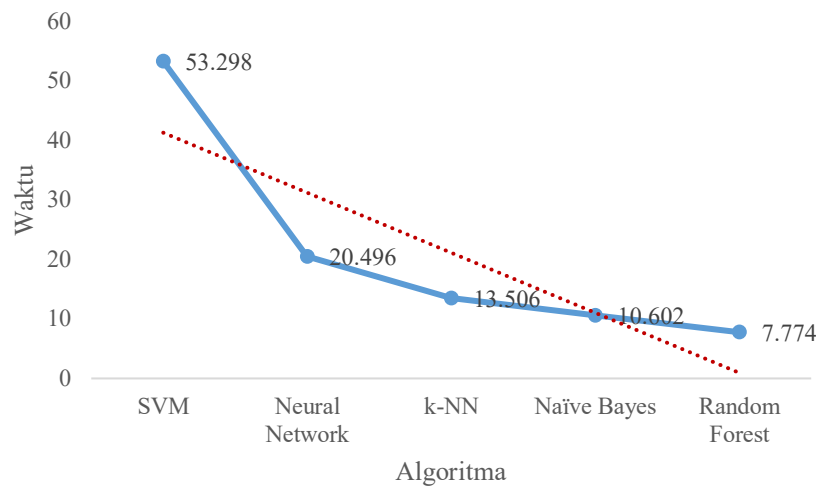
No	Algoritma	Akurasi	Precision	Recall
1	Naïve Bayes	0.636	0.645	0.636
2	SVM	0.818	0.820	0.818
3	Neural Network (NN)	0.844	0.845	0.844
4	Random Forest	0.648	0.647	0.648
5	k-NN	0.667	0.678	0.667

Hasil pengujian algoritma pada dataset citra daun tebu menampilkan kinerja yang berbeda-beda. Algoritma Neural Network (NN) secara konsisten menjadi yang terbaik. Hal ini karena secara implisit algoritma NN dapat mendeteksi dan memodelkan hubungan yang rumit dan nonlinier antara variabel input dan output, sesuatu yang sulit dilakukan oleh metode statistik atau algoritma tradisional lainnya. Sedangkan algoritma Naive Bayes menjadi yang terburuk. Nilai rata-rata algoritma berdasarkan akurasi, precision, dan recall ditampilkan pada Gambar 4.



Gambar 4. Perbandingan kinerja rata-rata

Gambar 4 menunjukkan bahwa nilai rata-rata algoritma NN secara konsisten menunjukkan yang terbaik disusul SVM. Sedangkan algoritma dengan nilai rata-rata terburuk yaitu Naive Bayes. Selanjutnya kami melakukan perbandingan waktu komputasi yang dibutuhkan algoritma dalam membangun model yang ditampilkan pada Gambar 5.



Gambar 5. Waktu komputasi algoritma

Hasil pengujian kinerja waktu komputasi, algoritma Random Forest menunjukkan waktu terbaik disusul algoritma Naive Bayes. Sedangkan algoritma SVM menunjukkan kinerja waktu terburuk. Meskipun secara akurasi algoritma Random Forest buruk, namun dari kecepatan komputasi menjadi yang tercepat.

KESIMPULAN DAN SARAN

Berdasarkan tahapan penelitian yang dirancang, diperoleh beberapa poin penting yaitu (a) Identifikasi penyakit daun tebu berbasis citra digital merupakan solusi efektif untuk mengatasi keterbatasan metode manual dalam mendeteksi penyakit tanaman di lapangan. (b) Perbandingan berbagai algoritma klasifikasi seperti Naive Bayes (NB), Neural Network (NN), k-Nearest Neighbors (k-NN), SVM, dan Random Forest memberikan pemahaman yang komprehensif mengenai trade-off antara akurasi dan efisiensi waktu komputasi. (c) Algoritma NN unggul dalam akurasi karena kemampuannya mengekstraksi fitur visual secara otomatis, namun memiliki kompleksitas komputasi tinggi. Sebaliknya, algoritma seperti Random Forest cenderung lebih ringan dan cepat, sehingga cocok untuk implementasi lapangan berbasis perangkat terbatas. (d) Penelitian ini menegaskan pentingnya evaluasi multi-aspek (akurasi, precision, recall, dan waktu) untuk menentukan model terbaik, bukan hanya berdasarkan akurasi semata.

Diharapkan dapat menjadi dasar pengembangan sistem deteksi otomatis penyakit tebu berbasis kamera atau drone, guna mendukung digitalisasi dan efisiensi sektor pertanian tebu di Indonesia.

UCAPAN TERIMA KASIH

Kami mengucapkan terimakasih kepada seluruh pihak terutama kepada Universitas Aisyah Pringsewu yang telah mendukung penuh penelitian ini. Kami juga ucapkan terimakasih kepada dosen Fakultas Teknologi dan Informatika dan petani Tebu di Kabupaten Pringsewu yang telah membantu memberikan masukan dan saran serta validasi data citra.

DAFTAR PUSTAKA

- Akkaya, E., & Turgay, S. (2024). Unveiling the Power: A Comparative Analysis of Data Mining Tools through Decision Tree Classification on the Bank Marketing Dataset. *Wseas Transactions on Computers*, 23, 95–105. <https://doi.org/10.37394/23205.2024.23.9>
- Chen, J.-L., Sun, S.-R., Wang, Z.-Q., Fu, H.-Y., Xu, H.-Y., Chang, H.-L., Gao, S.-J., & Wang, Q.-N. (2025). Identification of Causal Agents of Rust of *Saccharum* spp. and Assessment of Resistance to Brown Rust in *Erianthus arundinaceus* Clones and Their Offspring. In *Plants* (Vol. 14, Issue 8). <https://doi.org/10.3390/plants14081221>
- Hidayah, N. (2020). Peluang Pengembangan Pengendalian Penyakit Luka Api pada Tebu di Indonesia. *Buletin Tanaman Tembakau, Serat & Minyak Industri*, 12(2), 94–108. <https://doi.org/10.21082/btsm.v12n2.2020.94-108>
- Huang, J., Khan, M. T., Perecin, D., Coelho, S. T., & Zhang, M. (2020). Sugarcane for bioethanol production: Potential of bagasse in Chinese perspective. *Renewable and Sustainable Energy Reviews*, 133, 110296. <https://doi.org/https://doi.org/10.1016/j.rser.2020.110296>
- Kumar, M., Upadhyay, S. N., & Mishra, P. K. (2022). Pyrolysis of Sugarcane (*Saccharum officinarum* L.) Leaves and Characterization of Products. *ACS Omega*, 7(32), 28052–28064. <https://doi.org/10.1021/acsomega.2c02076>
- Li, X., Li, X., Zhang, M., Dong, Q., Zhang, G., Wang, Z., & Wei, P. (2024). SugarcaneGAN: A novel dataset generating approach for sugarcane leaf diseases based on lightweight hybrid CNN-Transformer network. *Computers and Electronics in Agriculture*, 219, 108762. <https://doi.org/https://doi.org/10.1016/j.compag.2024.108762>
- Srivastava, S., Kumar, P., Mohd, N., Singh, A., & Gill, F. S. (2020). A Novel Deep Learning Framework Approach for Sugarcane Disease Detection. *SN Computer Science*, 1(2), 87. <https://doi.org/10.1007/s42979-020-0094-9>
- Thite, S., Suryawanshi, Y., Patil, K., & Chumchu, P. (2024). Sugarcane leaf dataset: A dataset for disease detection and classification for machine learning applications. *Data in Brief*, 53, 110268. <https://doi.org/https://doi.org/10.1016/j.dib.2024.110268>