

Klasifikasi Penyakit Daun pada Tebu dengan Pendekatan Algoritma *K-Nearest Neighbors*, *Multilayer Perceptron* dan *Support Vector Machine*

Asep Syaputra

Program Studi Teknik Informatika, Institut Teknologi Pagar Alam
Jl. Simpang Bacang No.43, Karang Dalo, Dempo Tengah, Kota Pagar Alam, Sumatera Selatan, Indonesia
Email : asepsyaputra68@itpa.ac.id

ABSTRACT

Sugarcane is a vital crop in Indonesia, serving as the primary raw material for sugar production. Unfortunately, leaf diseases in sugarcane often pose a serious threat, potentially causing significant economic losses. These diseases are typically characterized by leaf morphological changes, making early detection and accurate classification essential to prevent further spread. This study compares three algorithms for identifying sugarcane leaf diseases: K-Nearest Neighbors (KNN), Multilayer Perceptron (MLP), and Support Vector Machine (SVM). Each algorithm employs a different approach to recognize patterns and disease characteristics: SVM separates data by identifying the optimal hyperplane, KNN classifies based on the proximity of data to training data, while MLP, as an artificial neural network, can recognize more complex patterns. The deep learning model VGG16 was utilized for feature extraction from sugarcane leaf images to enhance classification accuracy. The dataset used comprises 8,200 images of sugarcane leaves, categorized into four classes: 2,050 images of Cercospora spot gray, 2,050 of common rust, 2,050 of northern blight, and 2,050 of healthy leaves. Each category was further divided into training and testing datasets in an 80:20 ratio, with 6,560 images for training and 1,640 images for testing. The results indicate that the MLP algorithm achieved the best performance, with accuracy, precision, and recall values of 97.4%. This establishes MLP as the most effective choice for classifying sugarcane leaf diseases.

Keywords : *K-Nearest Neighbors (KNN), Multilayer Perceptron (MLP), Support Vector Machine (SVM)*

ABSTRAK

Tebu adalah tanaman penting di Indonesia yang menjadi bahan baku utama untuk produksi gula. Sayangnya, penyakit pada daun tebu sering menjadi ancaman serius yang dapat menyebabkan kerugian ekonomi besar. Penyakit ini biasanya ditandai dengan perubahan bentuk pada daun, sehingga pendeteksian dini dan pengklasifikasian yang akurat sangat penting untuk mencegah penyebaran penyakit yang lebih luas. Dalam penelitian ini, tiga algoritma dibandingkan untuk mengidentifikasi penyakit daun tebu, yaitu K-Nearest Neighbors (KNN), Multilayer Perceptron (MLP), dan Support Vector Machine (SVM). Ketiga algoritma ini memiliki cara kerja yang berbeda dalam mengenali pola dan karakteristik penyakit: SVM memisahkan data dengan mencari garis pemisah terbaik, KNN mengklasifikasikan berdasarkan kedekatan data dengan data latih, sementara MLP, sebagai jaringan saraf tiruan, mampu mengenali pola yang lebih kompleks. Untuk meningkatkan akurasi pengklasifikasian, model deep learning VGG16 digunakan untuk ekstraksi fitur gambar daun. Dataset yang digunakan mencakup 8.200 citra daun tebu yang diklasifikasikan ke dalam empat kategori: 2.050 citra bercak abu-abu (cercospora spot gray), 2.050 karat biasa (common rust), 2.050 hawar utara (northern blight), dan 2.050 citra daun sehat. Setiap kategori selanjutnya dibagi menjadi data pelatihan dan data pengujian dengan proporsi 80% untuk pelatihan (6.560 citra) dan 20% untuk pengujian (1.640 citra). Hasil penelitian menunjukkan bahwa algoritma MLP memberikan kinerja terbaik dengan nilai akurasi, presisi, dan recall sebesar 97,4%. Hal ini menjadikan MLP sebagai pilihan paling efektif untuk mengklasifikasikan penyakit daun tebu.

Kata Kunci : *K-Nearest Neighbors (KNN), Multilayer Perceptron (MLP), Support Vector Machine (SVM)*

1. Pendahuluan

Indonesia merupakan negara agraris, sektor pertanian menjadi mata pencarian utama bagi sebagian besar penduduknya (Ayun et al., 2020). Posisi geografis Indonesia yang terletak di kawasan tropis memberikan iklim yang mendukung pertumbuhan berbagai jenis tanaman, kondisi tanah yang subur, dan keberagaman sumber daya alam yang mendukung pengembangan

komoditas pertanian, termasuk tebu. Tebu sebagai salah satu komoditas penting memiliki peranan strategis dalam industri gula, yang menjadi bahan pokok masyarakat Indonesia. Pemanfaatan sumber daya agrikultur secara optimal merupakan kunci utama untuk meningkatkan produktivitas pertanian (Wahditiya et al., 2024). Oleh karena itu, pengelolaan dan alokasi terbatasnya sumber daya harus terapkan secara efisien agar bisa mencapai hasil maksimal. Berbagai faktor, seperti kualitas tanah,

teknologi yang diterapkan, dan keberlanjutan praktik pertanian, sangat memengaruhi hasil yang diperoleh. Selain sektor pertanian, Indonesia kaya akan sumber daya alam lainnya, contohnya sumber daya pada sector kehutanan, sumber daya ini menjadi komoditas utama yang dimanfaatkan dalam sektor perindustrian. Hal ini menunjukkan pentingnya pemanfaatan sumber daya alam secara berkelanjutan, yang tidak hanya mendukung pertumbuhan ekonomi, tetapi juga menjaga keseimbangan ekologis dan keberlanjutan lingkungan hidup di masa depan (Suwarno, 2024). Selain itu, sumber daya kehutanan juga berperan dalam memenuhi kebutuhan masyarakat, termasuk dalam mendukung produksi pangan seperti tanaman tebu (Kasmaniar et al., 2023).

Tebu adalah salah satu komoditas pertanian penting di Indonesia yang berperan sebagai bahan baku utama dalam industri gula, serta memiliki potensi besar dalam produksi bioetanol dan pakan ternak. Namun, meskipun memiliki nilai ekonomis tinggi, tanaman tebu rentan terhadap serangan penyakit daun yang dapat menyebabkan kerugian yang signifikan. Penyakit-penyakit ini umumnya dimulai dengan perubahan morfologi pada daun yang menjadi indikasi adanya gangguan pada tanaman. Deteksi dini dan klasifikasi penyakit secara akurat sangat penting untuk meminimalkan penyebaran penyakit tersebut dan mengurangi dampak kerugian pada hasil pertanian (Kasmaniar et al., 2023).

Pengidentifikasi penyakit pada tanaman tebu umumnya dilakukan oleh ahli pertanian, namun pendekatan ini dapat menghambat efisiensi pengelolaan kesehatan tanaman, khususnya bagi petani yang berada di daerah terpencil. Biaya konsultasi yang tinggi serta jarak yang jauh dari pusat layanan pertanian menjadi tantangan tersendiri yang dapat memperlambat respons terhadap serangan penyakit (Wardiman et al., 2024). Untuk mengatasi permasalahan ini, pemanfaatan teknologi komputer menawarkan solusi alternatif yang lebih efisien. Sistem berbasis komputer dapat menggantikan peran ahli dalam mendeteksi dan mengklasifikasikan penyakit tanaman menggunakan metode pengolahan citra dan algoritma klasifikasi dalam data mining. Dengan menggunakan sistem ini, petani dapat melakukan diagnosis penyakit secara lebih cepat, tepat, dan dengan biaya yang lebih terjangkau (Ningrum et al., 2024). Hal ini dapat meningkatkan efisiensi dalam penanganan penyakit dan pada gilirannya meningkatkan produktivitas tanaman tebu. Salah satu pendekatan yang efektif adalah dengan memanfaatkan model deep learning, seperti VGG16, yang mampu meningkatkan tingkat akurasi dalam pengenalan pola penyakit pada daun tanaman. Model ini berkontribusi pada sistem deteksi penyakit yang lebih presisi dan otomatis (Setiono, 2024).

Perkembangan pesat dalam teknologi komputasi dan peningkatan kapasitas pemrosesan komputer telah menciptakan peluang baru dalam otomatisasi berbagai proses, termasuk dalam bidang pertanian. Saat ini, komputer mampu melakukan pengenalan objek secara

otomatis melalui pemrosesan informasi yang diperoleh dari citra (Adha, 2020). Pengolahan citra, yang memungkinkan ekstraksi fitur dari gambar tanaman, memiliki peranan penting dalam mendeteksi, mengidentifikasi, dan mengklasifikasikan penyakit pada tanaman, termasuk penyakit yang menyerang daun tebu. Teknologi ini memungkinkan deteksi yang lebih cepat dan akurat, membantu petani dalam memitigasi penyebaran penyakit secara lebih efisien. Selain itu, dengan meningkatnya akurasi model-model pembelajaran mesin, seperti deep learning, kemampuan untuk mengidentifikasi penyakit tanaman secara otomatis dapat memberikan solusi yang lebih terjangkau dan dapat diakses oleh petani, terutama yang berada di daerah terpencil (Harjanti & Himawan, 2021). Data mining merupakan proses untuk mengekstraksi informasi pada kumpulan data besar, yang sangat berguna dalam mempercepat dan mempermudah identifikasi penyakit dengan mengkategorikannya berdasarkan label. Klasifikasi merupakan salah satu teknik dalam data mining yang sering digunakan, teknik ini memungkinkan pengelompokan penyakit tanaman berdasarkan ciri-ciri yang terdeteksi pada citra daun (Ikhwan & Aslami, 2020). Dengan menerapkan algoritma klasifikasi yang sesuai, seperti K-Nearest Neighbors (KNN), Multilayer Perceptron (MLP) dan Support Vector Machine (SVM) sistem komputer dapat mendukung para petani dalam mendeteksi penyakit pada tahap awal dan memungkinkan pengambilan tindakan preventif yang lebih cepat dan efisien, tanpa ketergantungan langsung pada ahli pertanian.

Berbagai penelitian terdahulu yang memanfaatkan teknik penambangan data dan pengolahan gambar untuk klasifikasi telah menghasilkan temuan yang signifikan. Penelitian yang sudah dilakukan oleh Putra et al. mengungkapkan bahwa metode Convolutional Neural Network (CNN) pada arsitektur ResNet 50 dan optimizer Adam berhasil mencapai akurasi 98,4% dalam mengklasifikasikan penyakit pada daun (Syahputra, 2023). Sementara itu, studi dari Hasan et al. memanfaatkan model CNN VGG16 sebagai alat mendeteksi penyakit pada daun tanaman anggur, yang menghasilkan akurasi pengujian sebesar 97,25% (Inaya et al., 2024). Di sisi lain, penelitian Kusuma et al. melakukan studi perbandingan algoritma Support Vector Machine (SVM) dan Naïve Bayes (NB) pada klasifikasi citra ras kucing, menunjukkan bahwa SVM memiliki performa tertinggi dengan 88,4% akurasi (Jinan et al., 2024). Sementara itu, penelitian yang dilakukan oleh Gunawan dan Putra mengevaluasi beberapa metode, yaitu Multi-Class SVM, Backpropagation Neural Network (NN), K-Nearest Neighbors (K-NN), dan Naïve Bayes (NB), untuk mengidentifikasi American Sign Language (Syahputra, 2023). Hasilnya menunjukkan bahwa Multi-Class SVM 99,3% mencapai akurasi tertinggi, diikuti oleh nilai 98,28% Backpropagation NN, 97,7% K-NN, dan 95,98% untuk NB.

Penelitian ini bertujuan pada membandingkan efektivitas algoritma K-Nearest Neighbors (KNN), Multilayer Perceptron (MLP), dan Support Vector

Machine (SVM) dalam mengklasifikasikan penyakit pada daun tebu. Proses ini menggunakan model deep learning VGG16 untuk mengekstrak fitur citra. Penelitian yang dilakukan oleh Veni dan Manjula menunjukkan bahwa model deep learning VGG16 terbukti dapat meningkatkan tingkat akurasi klasifikasi secara signifikan (Rahman et al., 2024). Selain itu, studi yang dilakukan oleh Prasetyo et al., yang membandingkan berbagai rangkaian deep learning seperti Mobile, Net V1 ResNet50, Xception, dan VGG16, juga menemukan bahwa model VGG-16 menghasilkan akurasi tertinggi dibandingkan dengan arsitektur lainnya (Akbar & Faisal, 2024).

2. Metode Penelitian

Tahapan penelitian dalam membandingkan algoritma KNN (K-Nearest Neighbors), MLP (Multilayer Perceptron), dan SVM (Support Vector Machine) untuk klasifikasi penyakit pada daun tebu umumnya mencakup beberapa langkah utama, yaitu: pengumpulan dataset gambar daun tebu, pra-pemrosesan data untuk meningkatkan kualitas gambar, ekstraksi fitur dari dataset, pelatihan model menggunakan algoritma yang dibandingkan, dan evaluasi kinerja model berdasarkan akurasi klasifikasi. Langkah-langkah ini dapat dilihat pada Gambar 1, yang menjelaskan proses klasifikasi penyakit daun pada tebu secara keseluruhan.



Gambar 1. Langkah-langkah dalam mengklasifikasikan penyakit daun pada tebu

2.1. Pengumpulan Data

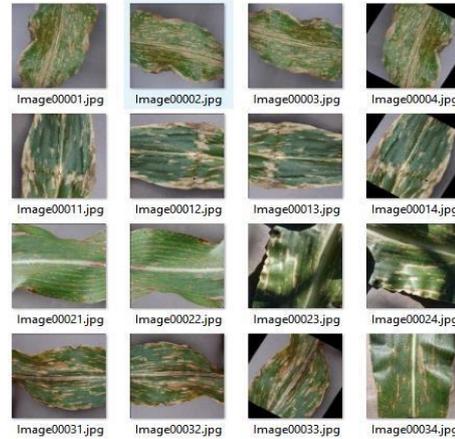
Penelitian ini menggunakan data berupa gambar daun tebu yang menunjukkan berbagai jenis penyakit, yang diperoleh dari sumber data sekunder yang dapat diakses melalui situs Kaggle. Data ini dipilih karena mencakup kategori penyakit daun yang sesuai dengan karakteristik tanaman jagung, sehingga relevan dengan fokus penelitian ini.

Dataset gambar ini terdiri dari 8.200 citra daun tebu yang dikelompokkan ke dalam 4 kategori: 2.050 citra cercospora spot gray, 2.050 common rust, 2.050 northern blight, dan 2.050 citra daun sehat. Setiap kategori gambar selanjutnya dipisahkan menjadi data pelatihan serta data pengujian dengan perbandingan 80% (data pelatihan 6.560) dan 20% (data pengujian 1.640). Pembagian ini dilakukan secara proporsional untuk keempat kategori tersebut agar setiap kelas memiliki

representasi yang seimbang dalam pelatihan dan pengujian model. Gambar 2 menunjukkan contoh dataset citra penyakit daun tebu untuk masing-masing kategori.

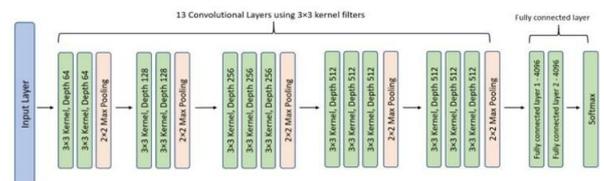
2.2. Preprocessing Data

Selanjutnya, dilakukan penyesuaian dataset sehingga

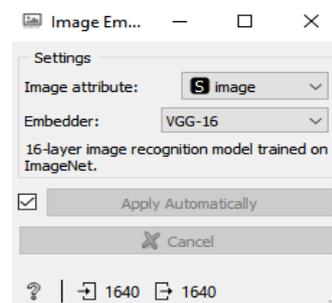


Gambar 2. Dataset penyakit daun tebu (citra)

dapat diolah oleh algoritma. Langkah-langkah dalam tahap praproses ini mencakup penyeragaman ukuran citra yang akan dimasukkan, dalam hal ini dengan citra dengan size 150*150 piksel. Selanjutnya, model deep learning VGG16 diterapkan untuk ekstraksi fitur citra. Proses-proses ini terlihat pada Gambar 3 dan Gambar 4, yang menggambarkan tahapan praproses dan ekstraksi fitur yang dilakukan untuk meningkatkan akurasi pada klasifikasi penyakit daun pada tebu.



Gambar 3. Struktur Arsitektur VGG16



Gambar 4. Model jaringan saraf VGG 16 atau Deep learning

2.3. Klasifikasi Citra

Tahap selanjutnya adalah mengklasifikasikan gambar penyakit pada daun tebu menggunakan algoritma KNN, MLP, dan SVM. Dalam tahap ini, model jaringan saraf dalam VGG16 diterapkan untuk mengekstrak fitur dari citra, dengan tujuan meningkatkan akurasi dalam mengidentifikasi jenis penyakit.

2.3.1 K-Nearest Neighbours

K-Nearest Neighbors (KNN) merupakan algoritma berbasis data yang sederhana dan tidak bergantung pada asumsi distribusi parametrik. Kelebihan utama dari KNN adalah tidak membutuhkan asumsi mengenai distribusi data yang mendasari, yang membuatnya cocok untuk aplikasi di dunia nyata, di mana banyak data tidak memenuhi asumsi matematis yang diperlukan oleh algoritma parametrik. Pada model KNN, tidak ada kebutuhan untuk pelatihan data eksplisit; seluruh dataset digunakan untuk mengklasifikasikan data uji, yang dapat mempercepat proses pelatihan, namun meningkatkan biaya dan waktu pengujian (Sujaini, 2019).

Namun, model ini juga memiliki beberapa keterbatasan. Dalam skenario terburuk, KNN bisa sangat memakan waktu, karena perlu memindai seluruh dataset untuk setiap data uji, serta memerlukan memori lebih untuk menyimpan data pelatihan. Penghitungan jarak antar titik data dalam KNN biasanya menggunakan metode Euclidean Distance (ED), yang digunakan sebagai alat ukur jarak antara dua titik pada ruang berdimensi-n

$$(x_1, x_2) = \sqrt{\sum(x_{1i} - x_{2i})^2} \quad (1)$$

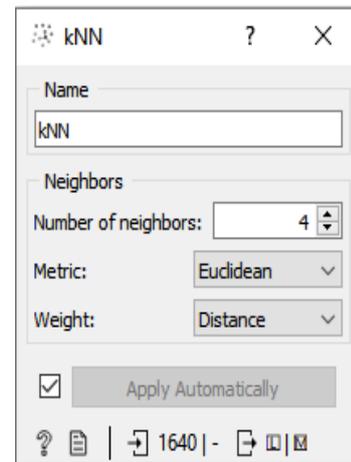
Pada KNN, jarak Euclidean (ED) digunakan untuk mengukur seberapa dekat dua titik data dalam ruang n-dimensi. Dalam hal ini, $D(x_1, x_2)$ merujuk pada jarak ED dari dua titik data, yaitu x_1 dan x_2 . Setiap titik data x_1 dan x_2 memiliki nilai atribut yang dihitung dengan membandingkan nilai dari setiap atribut i dalam kedua titik data, yaitu x_{1i} dan x_{2i} .

Setelah menghitung jarak ED, nilai k (jumlah tetangga terdekat) digunakan untuk menentukan kategori atau label dari data pengujian berdasarkan mayoritas kategori dari tetangga-tetangga terdekat. Sebagai contoh, apabila sebuah titik data baru dibandingkan dengan beberapa titik data tetangga, dan tiga dari tetangga tersebut berasal dari kategori A, sedangkan dua lainnya berasal dari kategori B, maka titik data baru tersebut akan dikategorikan ke dalam kategori A karena mayoritas tetangga terdekatnya berasal dari kategori tersebut.

Proses ini menunjukkan bagaimana KNN bekerja dengan memilih k tetangga terdekat untuk menentukan label dari data yang belum diketahui kategorinya. Dengan demikian, KNN merupakan algoritma yang sangat bergantung pada kedekatan data, yang mempengaruhi akurasi klasifikasi berdasarkan pemilihan nilai k yang tepat. Gambar 5 menggambarkan ilustrasi perhitungan jarak Euclidean (ED) antara titik data baru dan tetangga-tetangganya, sedangkan Gambar 6 menunjukkan penerapan model K-Nearest Neighbors (KNN) dalam klasifikasi gambar penyakit daun tebu.



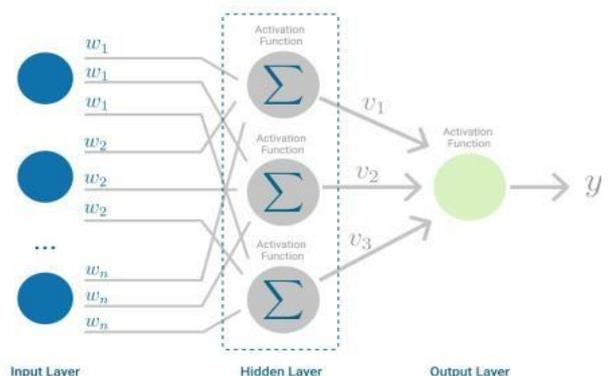
Gambar 5. Ilustrasi perhitungan jarak Euclidean (ED) pada titik data baru dengan *neighbours*



Gambar 6. Penerapan model *K-Nearest Neighbors* (KNN) dalam klasifikasi gambar penyakit daun tebu

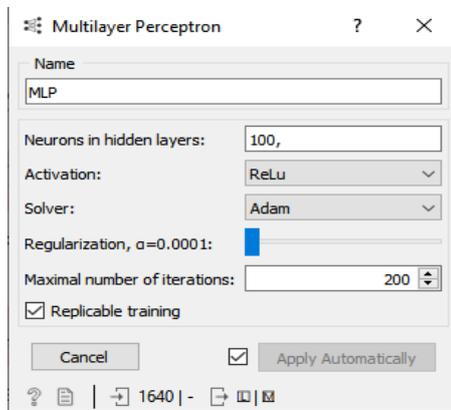
2.3.2 Perceptron Multilayer

MLP (*Perceptron Multilayer*) yaitu jenis jaringan saraf tiruan yang menggunakan arsitektur feedforward, yang terdiri dari 3 lapisan node, lapisan input, lapisan tersembunyi, serta lapisan output (Sinaga, 2020), seperti yang ditunjukkan pada Gambar 8. Setiap node dalam MLP, kecuali pada lapisan input, berperan sebagai neuron yang mengaplikasikan fungsi aktivasi non-linier. Selama proses pelatihan, MLP menerapkan teknik backpropagation. MLP berbeda dari linier perceptron karena memiliki beberapa lapisan dan menggunakan fungsi aktivasi nonlinier, memungkinkan jaringan ini memisahkan Data yang tidak dapat dibedakan secara linier. Proses ini dimulai dengan mengalikan input dengan bobot awal yang sudah ditentukan secara tertimbang, lalu menerapkannya pada fungsi aktivasi. Karena aliran informasi mengalir secara satu arah dari lapisan input menuju lapisan output, kategori algoritma feedforward termasuk pada MLP. Kombinasi linier pada jaringan MLP diteruskan ke lapisan selanjutnya untuk menjalani proses lebih lanjut. Gambar 7 memperlihatkan struktur utama MLP, yang terdiri dari tiga lapisan utama: lapisan input, lapisan tersembunyi (*hidden layer*), dan lapisan output.



Gambar 7. MLP memiliki tiga lapisan utama, lapisan input, lapisan hidden, serta lapisan output.

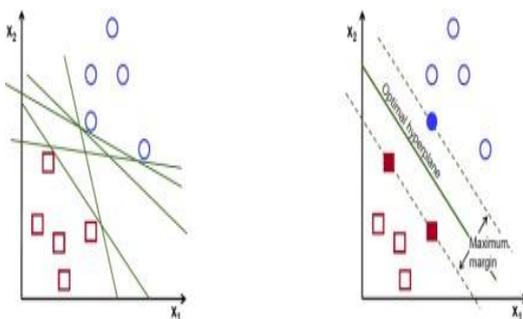
Penelitian ini menggunakan Arsitektur MLP yang diimplementasikan fungsi untuk aktivasi ReLU dan algoritma optimisasi Adam, yang jumlah 100 neuron pada lapisan tersembunyi. Model ini dapat dilihat pada Gambar 8.



Gambar 8. Klasifikasi citra penyakit daun tebu pada penerapan model *Perceptron Multilayer (MLP)*

2.3.3 SVM (Support Vector Machine)

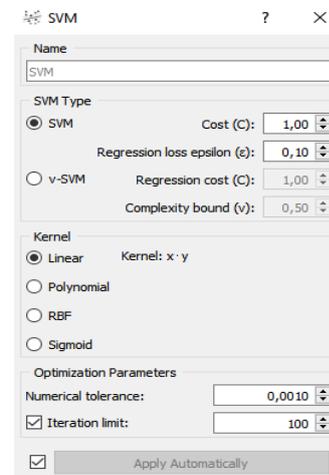
SVM (Support Vector Machine) merupakan pengklasifikasi diskriminatif yang menggunakan hyperplane untuk data dipisahkan ke dalam kelas-kelas yang berbeda. Algoritma ini berfokus pada pencarian hyperplane optimal yang memaksimalkan margin pemisah antara data dari kelas yang berbeda, dengan tujuan mengklasifikasikan data baru berdasarkan data pelatihan yang sudah dilabeli (Arumnisa & Wijayanto, 2023). Dalam kasus dua dimensi, hyperplane berupa sebuah garis pada bidang yang dibagi menjadi dua area, dimana setiap kelas, datanya berada pada sisi yang berlawanan dari garis pemisah tersebut.



Gambar 9. Hyperplane optimal ditentukan dengan cara jarak yang dimaksimalkan atau margin antara titik data dari kelas berbeda.

Seperti pada Gambar 9, terdapat berbagai kemungkinan hyperplane yang dapat dipilih untuk memisahkan dua kelompok titik data. Tujuan utama SVM adalah mencari hyperplane dengan margin terbesar, yaitu maksimum jarak antara dari kedua kelas yang memiliki titik data, untuk memastikan pemisahan yang optimal dan meningkatkan kemampuan generalisasi model. Memaksimalkan margin ini memberikan keuntungan dalam meningkatkan kepercayaan klasifikasi terhadap data baru. Hyperplane

bertindak sebagai batas pemisah yang mengklasifikasikan titik data ke dalam kelas yang berbeda, tergantung pada posisi relatifnya terhadap hyperplane tersebut. Dimensi hyperplane ditentukan oleh jumlah fitur dalam data. Model SVM yang digunakan terlihat pada Gambar 10.



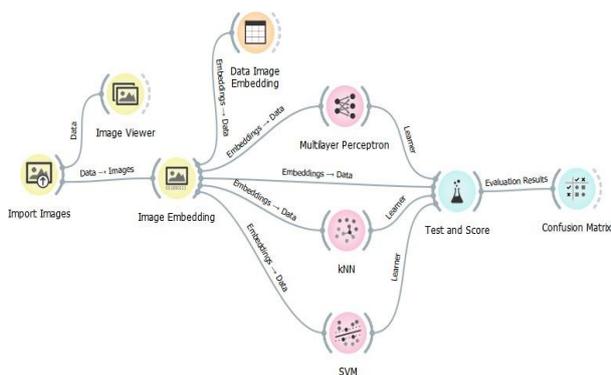
Gambar 10. Struktur klasifikasi untuk identifikasi penyakit pada daun tanaman tebu

2.3.4 Evaluasi Hasil dan Pengujian

Langkah berikutnya adalah melakukan evaluasi untuk mengukur tingkat akurasi tertinggi dalam klasifikasi gambar penyakit pada daun tanaman tebu. Proses evaluasi ini akan dipresentasikan dengan menggunakan Confusion Matrix, yang menggambarkan hasil prediksi dibandingkan dengan nilai aktual untuk mengevaluasi tingkat akurasi dari setiap metode yang diterapkan. Confusion Matrix berfungsi untuk mengidentifikasi jumlah prediksi yang benar dan salah, selain itu, untuk perhitungan evaluasi pada metric seperti akurasi, presisi, F1 score dan recall, yang menggambarkan lebih lengkap mengenai kinerja model.

3. Pembahasan

Proses pengklasifikasian dengan memanfaatkan aplikasi Orange Data Mining, dimana data dibagi dengan proporsi 80% untuk pelatihan dan 20% untuk pengujian, dengan jumlah data yang disesuaikan untuk setiap kategori. Penelitian ini mengaplikasikan algoritma KNN, MLP, dan SVM, di mana sebelum digunakan dalam pelatihan dan pengujian model, data telah melalui tahapan preprocessing untuk memastikan kualitas yang optimal. Proses preprocessing ini meliputi pembersihan data, normalisasi, dan ekstraksi fitur yang diperlukan untuk meningkatkan akurasi model. Desain dari alur proses klasifikasi ini dapat dilihat pada Gambar 11, yang menggambarkan secara rinci tahapan dalam penelitian ini, digunakan tiga algoritma yang berbeda untuk melakukan klasifikasi., diharapkan dapat diperoleh perbandingan yang lebih komprehensif mengenai kinerja masing-masing metode dalam mengklasifikasikan citra penyakit daun pada tanaman tebu.



Gambar 11. Penerapan model *Support Vector Machine (SVM)* dalam klasifikasi citra penyakit daun tebu.

Proses pengklasifikasian menggunakan algoritma SVM (*Support Vector Machine*) dengan data pelatihan. Berdasarkan Gambar 12, terlihat 6033 prediksi yang akurat, dengan rincian 1375 untuk cercospora leaf spot gray, kemudian 1639 untuk common rust, dan 1637 untuk kondisi sehat, selanjutnya 1382 untuk northern leaf blight. Di sisi lain, terdapat 527 prediksi yang salah, yang terbagi menjadi 253 untuk cercospora leaf spot gray, 3 untuk common rust, 10 untuk kondisi sehat, dan 261 untuk northern leaf blight, dari keseluruhan berjumlah 6560 data pelatihan.

		Predicted				
		Cercospora Leaf Spot Gray	Common Rust	Healthy	Northern Leaf Blight	Σ
Actual	Cercospora Leaf Spot Gray	1375	1	4	260	1640
	Common Rust	0	1639	0	1	1640
	Healthy	3	0	1637	0	1640
	Northern Leaf Blight	250	2	6	1382	1640
Σ		1628	1642	1647	1643	6560

Gambar 12. Confusion Matrix algoritma SVM yang diterapkan pada data training

Pada klasifikasi menggunakan algoritma KNN, data pelatihan digunakan untuk proses evaluasi. Seperti yang terlihat pada Gambar 13, tercatat 6273 prediksi yang benar, terdiri dari 1430 untuk cercospora leaf spot gray, kemudian 1621 untuk common rust, selanjutnya 1637 untuk kondisi sehat, dan 1585 untuk northern leaf blight. Di sisi lain, terdapat 287 prediksi yang salah, dengan rincian 37 untuk cercospora leaf spot gray, kemudian 11 untuk common rust, 35 untuk kondisi sehat, dan 204 untuk northern leaf blight, dari total 6560 data pelatihan.

		Predicted				
		Cercospora Leaf Spot Gray	Common Rust	Healthy	Northern Leaf Blight	Σ
Actual	Cercospora Leaf Spot Gray	1430	3	16	191	1640
	Common Rust	5	1621	2	12	1640
	Healthy	2	0	1637	1	1640
	Northern Leaf Blight	30	8	17	1585	1640
Σ		1467	1632	1672	1789	6560

Gambar 13. Algoritma KNN pada Confusion Matrix yang diterapkan pada data training

Pada klasifikasi menggunakan algoritma MLP, data pelatihan diterapkan dalam proses analisis. Gambar 14 menunjukkan bahwa terdapat 6391 prediksi yang akurat, dengan rincian untuk cercospora leaf spot gray nilai 1544, untuk common rust 1633, untuk kondisi sehat 1635, dan untuk northern leaf blight 1579. Sementara itu, terdapat 169 prediksi yang salah, yang terdiri dari 60 untuk cercospora leaf spot gray, 4 untuk common rust, 11 untuk kondisi sehat, dan 94 untuk northern leaf blight, pada total data pelatihan sebanyak 6560.

		Predicted				
		Cercospora Leaf Spot Gray	Common Rust	Healthy	Northern Leaf Blight	Σ
Actual	Cercospora Leaf Spot Gray	1544	2	4	90	1640
	Common Rust	4	1633	0	3	1640
	Healthy	4	0	1635	1	1640
	Northern Leaf Blight	52	2	7	1579	1640
Σ		1604	1637	1646	1673	6560

Gambar 14. Algoritma MLP pada Confusion Matrix yang diterapkan pada data training

Hasil uji yang telah dilakukan pada data pelatihan dapat dilihat pada Tabel 1. Berdasarkan temuan tersebut, dapat disimpulkan untuk algoritma MLP menunjukkan performa terbaik pada akurasi, presisi, dan recall masing-masing mencapai 97,4%. Hal ini mengindikasikan untuk MLP memiliki kemampuan pada pengklasifikasian data dengan tingkat akurasi yang sangat tinggi dibandingkan dengan algoritma lainnya yang diuji. Keunggulan ini memperkuat potensi MLP sebagai model yang efektif dan dapat diandalkan untuk mendeteksi penyakit pada daun tanaman tebu, dengan tingkat kesalahan klasifikasi yang relatif rendah.

Tabel 1. Hasil uji klasifikasi dengan data training

Algoritma	Accuracy	Precision	Recall
SVM	92,0%	91,9%	92,0%
KKN	95,6%	95,8%	95,6%
MLP	97,4%	97,4%	97,4%

Proses klasifikasi menggunakan algoritma SVM dilakukan pada data pengujian. Berdasarkan Gambar 15, terdapat 1539 prediksi yang akurat, yang terbagi menjadi untuk cercospora leaf spot gray 363, untuk common rust 407, untuk kondisi sehat 406, dan untuk northern leaf blight 363. Di sisi lain, terdapat 101 prediksi yang salah, dengan rincian 45 untuk cercospora leaf spot gray, untuk common rust ada 3, untuk kondisi sehat ada 2, dan untuk northern leaf blight 51, dari total data pengujian sebanyak 1640. Hasil ini menunjukkan bahwa SVM mendapatkan akurasi yang tinggi dalam mengklasifikasikan data testing, meskipun terdapat beberapa kesalahan dalam prediksi. Evaluasi ini memberikan gambaran tentang keandalan SVM dalam mendeteksi penyakit daun, sekaligus menunjukkan area

dimana perbaikan dapat dilakukan untuk lebih mengoptimalkan akurasi model.

		Predicted				
		Cercospora Leaf Spot Gray	Common Rust	Healthy	Northern Leaf Blight	Σ
Actual	Cercospora Leaf Spot Gray	363	1	0	46	410
	Common Rust	1	407	0	2	410
	Healthy	1	0	406	3	410
	Northern Leaf Blight	43	2	2	363	410
Σ		408	410	408	414	1640

Gambar 15. Confusion Matrix SVM yang diterapkan pada testing data

Pada klasifikasi yang menggunakan algoritma KNN, proses pengujian dilakukan pada data uji. Gambar 16 memperlihatkan bahwa terdapat 1511 prediksi yang benar, dengan rincian untuk cercospora leaf spot gray 328, untuk common rust 396, untuk kondisi sehat 408, dan untuk northern leaf blight 379. Sementara itu, ada 129 prediksi yang salah, yang terbagi menjadi untuk cercospora leaf spot gray 17, untuk common rust 8, untuk kondisi sehat 14, dan untuk northern leaf blight 90, dari total data pengujian sebanyak 1640. Hasil ini mendapatkan bahwa algoritma KNN memiliki kinerja yang cukup baik pada pengklasifikasian data testing, meskipun terdapat sejumlah kesalahan prediksi. Evaluasi ini memberikan wawasan tentang efektivitas KNN, sekaligus area di mana akurasi model dapat ditingkatkan.

		Predicted				
		Cercospora Leaf Spot Gray	Common Rust	Healthy	Northern Leaf Blight	Σ
Actual	Cercospora Leaf Spot Gray	328	1	5	76	410
	Common Rust	1	396	0	13	410
	Healthy	1	0	408	1	410
	Northern Leaf Blight	15	7	9	379	410
Σ		345	404	422	469	1640

Gambar 16. Matriks algoritma KNN yang diterapkan pada data pengujian

Proses klasifikasi menggunakan algoritma MLP diterapkan pada data uji. Berdasarkan Gambar 17, tercatat 1549 prediksi yang akurat, dengan rincian untuk cercospora leaf spot gray 365, untuk common rust 405, untuk kondisi sehat 408, dan untuk northern leaf blight 371. Sebaliknya, terdapat prediksi yang salah 91, yang terdiri dari untuk cercospora leaf spot gray 31, untuk common rust 9, untuk kondisi sehat 11, dan untuk northern leaf blight 40, dari total data uji sebanyak 1640. Hasil ini menunjukkan bahwa MLP memiliki kinerja yang luar biasa dalam mengklasifikasikan data pengujian, dengan tingkat kesalahan yang relatif rendah. Evaluasi ini mengonfirmasi efektivitas MLP dalam mendeteksi penyakit daun pada tanaman, serta memberikan wawasan mengenai area yang masih perlu ditingkatkan untuk mencapai tingkat akurasi yang lebih optimal.

		Predicted				
		Cercospora Leaf Spot Gray	Common Rust	Healthy	Northern Leaf Blight	Σ
Actual	Cercospora Leaf Spot Gray	365	3	3	39	410
	Common Rust	4	405	0	1	410
	Healthy	2	0	408	0	410
	Northern Leaf Blight	25	6	8	371	410
Σ		396	414	419	411	1640

Gambar 17. Confusion Matriks dari algoritma MLP yang diterapkan pada data uji

Pada Tabel 2 adalah hasil pengujian, bisa disimpulkan untuk algoritma MLP mendapatkan performa terbaik, dengan akurasi 94,5%, presisi 94,4%, dan 94,5% recall. Temuan ini menunjukkan bahwa MLP memiliki kemampuan klasifikasi yang sangat akurat dan konsisten, lebih unggul dibandingkan dengan algoritma lainnya, menjadikannya pilihan yang sangat efektif untuk mendeteksi penyakit daun pada tanaman. Tingginya nilai akurasi, presisi, dan recall ini menunjukkan bahwa MLP berhasil meminimalisir kesalahan klasifikasi, sekaligus menunjukkan kemampuannya dalam mendeteksi hampir seluruh kasus penyakit daun dengan tepat.

Tabel 2. Hasil uji klasifikasi dengan testing data

Algoritma	Accuracy	Precision	Recall
SVM	93,8%	93,9%	93,8%
KKN	92,1%	92,6%	92,1%
MLP	94,5%	94,4%	94,5%

Hasil Perbandingan dari data uji adalah Pada data pelatihan, algoritma MLP menunjukkan performa terbaik dengan akurasi, presisi, dan recall masing-masing mencapai 97,4%. Tingginya performa ini mencerminkan kemampuan MLP untuk mengklasifikasikan data pelatihan secara hampir sempurna, dengan tingkat kesalahan klasifikasi yang sangat rendah. Keunggulan ini menegaskan bahwa MLP adalah algoritma yang sangat andal dalam mengenali pola pada data pelatihan, menjadikannya pilihan utama untuk tugas klasifikasi penyakit pada daun tebu.

Ketika diuji pada data pengujian, MLP kembali menunjukkan keunggulannya dengan akurasi sebesar 94,5%, presisi 94,4%, dan recall 94,5%. Meskipun terdapat sedikit penurunan dibandingkan hasil pada data pelatihan, performa MLP tetap lebih tinggi dibandingkan algoritma lainnya. Hal ini menunjukkan bahwa MLP memiliki kemampuan yang konsisten dalam mendeteksi penyakit daun tebu, bahkan pada data yang belum pernah dilihat sebelumnya. Penurunan yang relatif kecil ini juga mencerminkan kemampuan generalisasi yang baik dari model MLP.

3. Kesimpulan

Uji coba terhadap algoritma *KNN (K-Nearest Neighbors)*, *MLP (Multilayer Perceptron)*, dan *SVM (Support Vector Machine)* dilakukan dengan memanfaatkan aplikasi Orange pada Data Mining. Menggunakan Dataset yang berisi citra penyakit daun tebu yang diambil dari sumber data sekunder yang tersedia di platform Kaggle. Dataset ini terdiri dari 8200 gambar yang terbagi ke dalam empat kategori: 2050 citra cercospora leaf spot gray, 2050 citra common rust, 2050 citra northern leaf blight, dan 2050 citra kondisi sehat. Sebelum proses klasifikasi, dilakukan pra-pemrosesan untuk menyesuaikan ukuran gambar menjadi 150x150 piksel. Fitur citra kemudian diekstraksi dengan model deep learning VGG16. Hasil eksperimen menunjukkan jika algoritma Multilayer Perceptron (MLP) memberikan performa terbaik yang memiliki akurasi, presisi, dan recall sebesar 97,4%. Kesimpulannya, hasil ini menunjukkan bahwa MLP unggul dibandingkan dengan algoritma lainnya dalam hal keakuratan dan kemampuan klasifikasi, menjadikannya pilihan yang sangat baik untuk aplikasi klasifikasi citra penyakit tanaman. Selain itu, hasil ini menyoroti pentingnya tahapan preprocessing dan ekstraksi fitur dalam meningkatkan performa model dalam mendeteksi penyakit daun

Daftar Pustaka

- Adha, L. A. (2020). Digitalisasi Industri Dan Pengaruhnya Terhadap Ketenagakerjaan Dan Hubungan Kerja Di Indonesia. *Jurnal Kompilasi Hukum*, 5(2), 267–298.
- Akbar, I., & Faisal, M. (2024). Perbandingan Analisis Sentimen PLN Mobile: Machine Learning Vs. Deep Learning. *JOINTECS (Journal Of Information Technology And Computer Science)*, 8(1), 1–10.
- Arumnisaa, R. I., & Wijayanto, A. W. (2023). Comparison Of Ensemble Learning Method: Random Forest, Support Vector Machine, Adaboost For Classification Human Development Index (HDI). *Sistemasi: Jurnal Sistem Informasi*, 12(1), 206–218.
- Ayun, Q., Kurniawan, S., & Saputro, W. A. (2020). Perkembangan Konversi Lahan Pertanian Di Bagian Negara Agraris. *Vigor: Jurnal Ilmu Pertanian Tropika Dan Subtropika*, 5(2), 38–44.
- Harjanti, T. W., & Himawan, H. (2021). Teknologi Pengolahan Citra Digital Untuk Ekstraksi Ciri Pada Citra Daun Untuk Identifikasi Tumbuhan Obat. *Faktor Exacta*, 14(3), 150–159.
- Ikhwan, A., & Aslami, N. (2020). Implementasi Data Mining Untuk Manajemen Bantuan Sosial Menggunakan Algoritma K-Means. (*Jurti*) *Jurnal Teknologi Informasi*, 4(2), 208–217.
- Inaya, A. N., Rahma, A. U., Jannah, M., Arafah, L. R. K., Ishak, L. L., & Edy, M. R. (2024). Klasifikasi Citra Dengan Pendekatan Transfer Learning Pada Gambar Fauna Terbang. *Jurnal Mediatik*, 85–89.
- Jinan, A., Siregar, M., Rolanda, V., Suryani, D. F., & Muis, A. (2024). Comparing Neural Networks, Support Vector Machines, And Naïve Bayes Algorithms For Classifying Banana Types. *Journal Of Computer Networks, Architecture And High Performance Computing*, 6(1), 98–107.
- Kasmaniar, K., Yana, S., Nelly, N., Fitriliana, F., Susanti, S., Hanum, F., & Rahmatullah, A. (2023). Pengembangan Energi Terbarukan Biomassa Dari Sumber Pertanian, Perkebunan Dan Hasil Hutan: Kajian Pengembangan Dan Kendalanya. *Jurnal Serambi Engineering*, 8(1).
- Ningrum, B. N. T. C., Ni'mah, E. N., Arifin, M. P., & Dara, M. A. D. W. (2024). KLASIFIKASI DAN PENGENALAN POLA PENYAKIT CABAI DENGAN METODE CNN (Convolution Neural Network). *Seminar Nasional Teknologi & Sains*, 3(1), 125–132.
- Rahman, A. T., Setyanto, A., & Al Fatta, H. (2024). Klasifikasi Penyakit Daun Apel Menggunakan Arsitektur CNN Dengan Transfer Learning. *Jurnal SENOPATI: Sustainability, Ergonomics, Optimization, And Application Of Industrial Engineering*, 6(1).
- Setiono, M. (2024). Klasifikasi Penyakit Antraknosa Citra Cabai Rawit Dengan Metode Convolutional Neural Network (CNN). *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, 11(2).
- Sinaga, D. (2020). Jaringan Saraf Tiruan Infeksi Mata Dengan Menggunakan Metode Berarsitektur Multi Layer Perceptron. *Informasi Dan Teknologi Ilmiah (INTI)*, 7(2), 189–192.
- Sujaini, H. (2019). Klasifikasi Citra Alat Musik Tradisional Dengan Metode K-Nearest Neighbor, Random Forest, Dan Support Vector Machine. *Jurnal Sistem Informasi Bisnis*, 9(2), 185–191.
- Suwarno, R. N. (2024). Strategi Ketahanan Pangan Dari Basis Lokal: Integrasi Prinsip Permakultur Dalam Teknologi Pangan Yang Berkelanjutan. *Indonesian Journal Of Applied Science And Technology*, 5(2), 52–66.
- Syahputra, F. (2023). *Analisis Arsitektur Deep Learning Mobilenet Dalam Mengklasifikasi Hama Daun Jambu Madu*. Universitas Medan Area.
- Wahditiya, A. A., Kurniawan, A., Nendissa, J. I., Meyuliana, A., Yora, M., Jamilah, J., Ilham, D. J., Mufaidah, I., Alaydrus, A. Z. A., & Hidayati, F. (2024). *Teknologi Produksi Tanaman Pangan*. Yayasan Tri Edukasi Ilmiah.
- Wardiman, B., Fitriyani, E., Herlyani, S., Ashar, J. R., & Panga, N. J. (2024). *Pertanian Keberlanjutan*. TOHAR MEDIA.