

Klasifikasi Kualitas Tandan Buah Segar Kelapa Sawit Menggunakan Support Vector Machine dengan Kernel Radial Basis Function

Latifa Khoirani^{1*}, Suhardi²

^{1,2} Prodi Ilmu Komputer, Fakultas Sains dan Teknologi, Universitas Islam Negeri Sumatera Utara, Medan, Indonesia

*¹ khoiranilatifa@gmail.com, ² suhardi@uinsu.ac.id

<https://doi.org/10.56427/jcbd.v5i2.993>

INFO ARTIKEL

Histori Artikel

Diterima : 25 April 2026

Direvisi : 26 Mei 2026

Disetujui : 30 Mei 2026

Kata Kunci

Support Vector Machine

Klasifikasi

Kelapa Sawit

TBS

Machine Learning

Keywords

Support Vector Machine

Classification

Palm Oil

FFB

Machine Learning

ABSTRAK

Kualitas Tandan Buah Segar (TBS) kelapa sawit menentukan mutu minyak sawit yang dihasilkan. Klasifikasi kualitas TBS di lapangan masih dilakukan secara manual berdasarkan pengamatan visual yang subjektif dan tidak konsisten. Penelitian ini menerapkan metode Support Vector Machine (SVM) dengan kernel Radial Basis Function (RBF) untuk mengklasifikasikan kualitas TBS secara otomatis berbasis data numerik parameter laboratorium sebagai alternatif yang lebih praktis dan efisien. Data yang digunakan sebanyak 1.500 dataset hasil pemeriksaan laboratorium periode Februari 2025 dari PT XYZ, dengan parameter berat tandan, kandungan minyak, kadar air, dan brondol lepas. Preprocessing meliputi data cleaning, normalisasi Min-Max Scaling, dan encoding label, dengan pembagian data 80:20 menghasilkan 1.200 data training dan 300 data testing. Model dibangun dengan parameter $C = 1,0$ dan $\gamma = 1,0542$ yang dihitung secara dinamis dari distribusi data training, menggunakan strategi klasifikasi multikelas One-vs-One (OvO). Evaluasi model menggunakan confusion matrix dengan metrik akurasi, precision, recall, dan F1-score. Hasil penelitian menunjukkan model SVM mampu mengklasifikasikan kualitas TBS ke dalam tiga kategori mentah, matang, dan lewat matang dengan akurasi 95%. Penelitian ini berkontribusi dalam meningkatkan objektivitas, konsistensi, dan efisiensi proses klasifikasi kualitas TBS berbasis machine learning.

The quality of Fresh Fruit Bunches (FFB) of oil palm determines the quality of the palm oil produced. FFB quality classification in the field is still performed manually through visual inspection, which tends to be subjective and inconsistent. This study applies the Support Vector Machine (SVM) method with a Radial Basis Function (RBF) kernel to automatically classify FFB quality based on numerical laboratory parameter data as a more practical and efficient alternative. The dataset consists of 1,500 records from laboratory examinations conducted in February 2025 at PT XYZ, encompassing the parameters of bunch weight, oil content, moisture content, and loose fruitlets. Preprocessing stages included data cleaning, Min-Max Scaling normalization, and label encoding, with an 80:20 data split yielding 1,200 training samples and 300 testing samples. The model was constructed with parameters $C = 1.0$ and $\gamma = 1.0542$, calculated dynamically from the training data distribution, employing a One-vs-One (OvO) multiclass classification strategy. Model evaluation was conducted using a confusion matrix with the metrics of accuracy, precision, recall, and F1-score. The results demonstrate that the SVM model is capable of classifying FFB quality into three categories unripe, ripe, and overripe achieving an accuracy of 95%. This study contributes to enhancing the objectivity, consistency, and efficiency of machine learning-based FFB quality classification.



JCBD is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-nc-sa/4.0/).

1. Pendahuluan

Kelapa sawit merupakan salah satu komoditas perkebunan utama yang memiliki kontribusi besar terhadap perekonomian Indonesia, baik dari sisi ekspor maupun penyerapan tenaga kerja. Produksi minyak kelapa sawit mentah (Crude Palm Oil/CPO) sangat bergantung pada kualitas bahan baku yang digunakan, yaitu Tandan Buah Segar (TBS) [1]. Kualitas TBS menjadi faktor kunci dalam menentukan rendemen minyak serta kadar asam lemak bebas yang dihasilkan. Semakin baik kualitas TBS, maka semakin optimal pula hasil produksi minyak sawit [2]. Oleh karena itu, pengendalian kualitas TBS menjadi aspek penting dalam menjaga daya saing industri kelapa sawit di pasar global [3].

Namun demikian, proses penentuan kualitas TBS di lapangan masih dilakukan secara manual melalui pengamatan visual oleh petugas [4]. Penilaian ini umumnya didasarkan pada indikator seperti warna buah, jumlah brondol lepas, dan kondisi fisik

tandan. Metode ini memiliki kelemahan karena sangat bergantung pada pengalaman dan subjektivitas individu, sehingga sering menimbulkan perbedaan penilaian antar petugas [5]. Kondisi ini menyebabkan hasil klasifikasi yang tidak konsisten dan berpotensi menurunkan kualitas data yang digunakan dalam proses pengolahan di pabrik [6]. Permasalahan tersebut berdampak pada ketidaksesuaian antara kualitas TBS yang ditentukan di kebun dengan hasil pengujian di pabrik. Ketidaksesuaian ini dapat menyebabkan penurunan efisiensi produksi, peningkatan biaya operasional, serta menurunnya kualitas minyak yang dihasilkan. Selain itu, kesalahan dalam klasifikasi juga dapat menghambat proses pengambilan keputusan yang tepat dalam manajemen perkebunan [7]. Oleh karena itu, diperlukan suatu sistem yang mampu melakukan klasifikasi kualitas TBS secara lebih objektif, akurat, dan konsisten berbasis data [8].

Untuk mengatasi permasalahan tersebut, beberapa penelitian sebelumnya telah mengembangkan sistem klasifikasi kematangan TBS dengan berbagai pendekatan. Zolfagharnassab et al. mengklasifikasikan kematangan TBS menggunakan teknik pencitraan termal (thermal imaging). Sari et al. menerapkan probe optik untuk mengklasifikasikan tingkat kematangan TBS berdasarkan karakteristik cahaya yang diserap buah. Kurniawan et al. menggunakan arsitektur deep learning YOLOv5 untuk deteksi kematangan buah sawit berbasis citra. Meskipun pendekatan tersebut menunjukkan hasil yang baik, metode berbasis citra dan sensor memerlukan infrastruktur pengambilan gambar yang kompleks, sensitif terhadap kondisi pencahayaan lapangan, serta membutuhkan dataset citra dalam jumlah besar untuk proses pelatihan model.

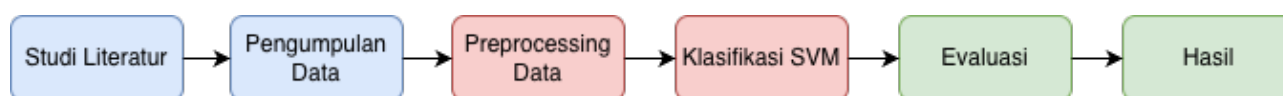
Di sisi lain, data hasil pemeriksaan laboratorium harian yang mencakup parameter numerik berat tandan, kandungan minyak, kadar air, dan brondol lepas tersedia secara rutin di setiap pabrik pengolahan kelapa sawit [9]. Namun, data tersebut belum dimanfaatkan secara optimal sebagai dasar sistem klasifikasi kualitas TBS berbasis machine learning [10]. Penelitian yang menggunakan Support Vector Machine (SVM) untuk berbagai tugas klasifikasi telah menunjukkan performa yang baik dibandingkan algoritma lain, namun belum terdapat penelitian yang secara spesifik menerapkan SVM pada data numerik parameter laboratorium untuk klasifikasi kualitas TBS. Kesenjangan inilah yang menjadi motivasi penelitian ini untuk membangun sistem klasifikasi yang praktis tanpa bergantung pada perangkat citra atau sensor khusus [11].

Metode SVM dipilih dalam penelitian ini karena kemampuannya menemukan hyperplane optimal yang memisahkan data multikelas dengan margin maksimum serta efektif menangani data berdimensi tinggi [12]. Penggunaan kernel Radial Basis Function (RBF) memungkinkan SVM menangani data yang tidak dapat dipisahkan secara linear [13], yang sesuai dengan karakteristik distribusi parameter numerik kualitas TBS. Berdasarkan keunggulan tersebut, penelitian ini bertujuan untuk menerapkan metode Support Vector Machine dengan kernel RBF dalam mengklasifikasikan kualitas TBS kelapa sawit ke dalam tiga kategori—mentah, matang, dan lewat matang—berdasarkan data numerik hasil pemeriksaan laboratorium dari PT XYZ. Diharapkan hasil penelitian ini dapat meningkatkan akurasi, objektivitas, dan efisiensi proses penilaian kualitas TBS serta mendukung pengembangan sistem klasifikasi otomatis di industri kelapa sawit [14].

2. Metodologi Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan data Tandan Buah Segar (TBS) kelapa sawit sebanyak 1.500 dataset yang diperoleh dari PT XYZ periode Februari 2025. Data tersebut terdiri dari beberapa parameter yang berpengaruh terhadap kualitas TBS, yaitu berat tandan (kg), kandungan minyak (%), kadar air (%), dan brondol lepas (%) [15]. Variabel target atau label adalah kualitas TBS yang dikategorikan ke dalam tiga kelas, yaitu mentah, matang, dan lewat matang. Implementasi dilakukan menggunakan Bahasa pemrograman python melalui platform google colab dengan library utama scikit-learn, pandas, numpy, matplotlib dan seaborn.

Proses labeling data dilakukan berdasarkan standar kriteria kualitas TBS yang ditetapkan oleh PT XYZ, mencakup parameter kematangan buah berdasarkan fraksi brondol lepas per tandan, rentang nilai kadar minyak, dan ambang batas kadar air. Klasifikasi label dilakukan oleh petugas laboratorium yang berpengalaman sesuai prosedur baku perusahaan, sehingga konsistensi antar petugas terjaga melalui pedoman kriteria yang terstandarisasi. Alur kerangka penelitian secara keseluruhan ditunjukkan pada Gambar 1.



Gambar 1. Kerangka penelitian

Studi Literatur

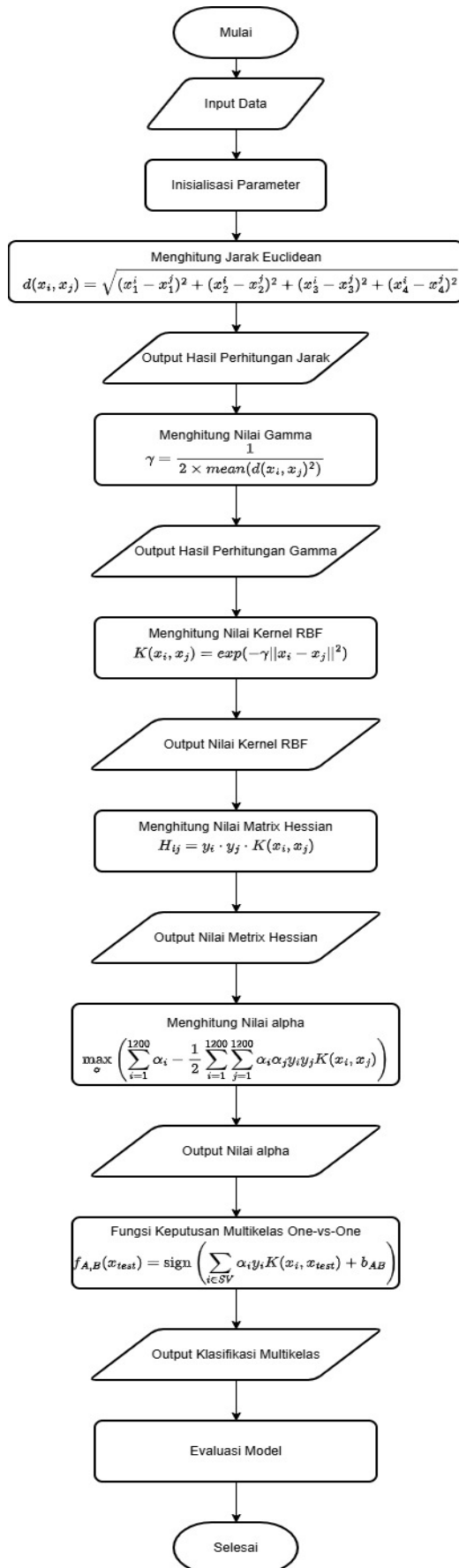
Tahapan penelitian dimulai dari studi literatur untuk memperoleh referensi dan teori pendukung yang relevan dengan topik penelitian. Selanjutnya dilakukan pengumpulan data dari sumber yang telah ditentukan. Tahapan pertama dalam penelitian ini yaitu studi literatur, yaitu pengumpulan sumber data pustaka dari buku, jurnal maupun karya ilmiah terdahulu, serta skripsi terdahulu yang dijadikan sebagai acuan data maupun model yang berhubungan dengan penelitian yang akan dilakukan. Setelah sumber data pustaka terkumpulkan, penulis mengkombinasikan beberapa model klasifikasi dari sumber pustaka terdahulu dengan rumusan masalah yang belum pernah dilakukan sebelumnya.

Pengumpulan Data

Selanjutnya, tahapan pengumpulan data penulis lakukan dengan mengumpulkan data TBS kelapa sawit yang ada di PT XYZ. Data yang digunakan dalam penelitian ini berjumlah 1.500 data hasil pemeriksaan laboratorium harian periode Februari 2025. Data TBS kelapa sawit yang digunakan memiliki fitur atau atribut yakni berat tandan, kandungan minyak, kadar air dan brondol lepas.

Preprocessing Data

Setelah data terkumpul, dilakukan preprocessing data yang meliputi data cleaning untuk mengatasi missing value dan data duplikat, normalisasi menggunakan metode Min-Max untuk mengubah rentang nilai setiap fitur ke skala 0 dan 1 agar tidak terjadi dominasi fitur berskala besar terhadap model, serta encoding label untuk mengubah kelas kualitas dari bentuk kategorikal menjadi numerik.



Gambar 2. Flowchart klasifikasi

Klasifikasi SVM

Setelah tahap preprocessing, data dibagi menjadi dua bagian yaitu data training sebesar 80% (1.200 data) dan data testing sebesar 20% (300 data) menggunakan teknik stratified split agar distribusi kelas tetap seimbang. Model klasifikasi dibangun menggunakan metode Support Vector Machine (SVM) dengan kernel Radial Basis Function (RBF) karena mampu menangani data non-linear dan memiliki performa yang baik dalam klasifikasi. Parameter utama yang digunakan adalah nilai C (cost) dan gamma (γ) yang diperoleh melalui proses perhitungan dan eksperimen untuk menghasilkan model yang optimal. Untuk klasifikasi multikelas digunakan strategi One-vs-One (OvO) [16].

Proses klasifikasi dimulai dengan menghitung jarak Euclidean antar data yang berfungsi sebagai dasar untuk menentukan tingkat kemiripan karakteristik antar data TBS, kemudian menentukan nilai gamma yang dihitung menggunakan rata-rata jarak Euclidean untuk mendapatkan parameter yang paling representative terhadap pola data TBS sehingga model dapat mengenali batasan tiap kelas dengan lebih detail, dilanjutkan dengan mencari nilai alfa (α) dan bias (b) untuk menemukan posisi bidang pemisah (hyperplane) paling optimal antara kelas mentah, matang dan lewat matang. Setelah itu, dilakukan proses klasifikasi multikelas menggunakan fungsi keputusan One-vs-One untuk menentukan kelas akhir dari data uji. Alur proses klasifikasi ini ditunjukkan pada Gambar 2.

Evaluasi Model

Evaluasi model dilakukan menggunakan confusion matrix untuk menghitung nilai akurasi, precision, recall, dan F1-score guna mengetahui performa model dalam mengklasifikasikan kualitas TBS secara keseluruhan. Setelah mendapatkan keseluruhan nilai evaluasi, nilai-nilai tersebut dimasukkan ke tahapan terakhir yaitu tahapan hasil untuk diambil kesimpulan dari hasil evaluasi model klasifikasi SVM tersebut.

3. Hasil dan Pembahasan

Penelitian ini menghasilkan model klasifikasi kualitas Tandan Buah Segar (TBS) kelapa sawit menggunakan metode Support Vector Machine (SVM) dengan kernel Radial Basis Function (RBF). Dataset yang digunakan berjumlah 1.500 data, dengan pembagian 1.200 data training dan 300 data testing menggunakan teknik stratified split. Model dilatih untuk mengklasifikasikan data ke dalam tiga kelas, yaitu mentah, matang, dan lewat matang berdasarkan parameter berat tandan, kandungan minyak, kadar air, dan brondol lepas.

Hasil Preprocessing Data

Tahap preprocessing dilakukan untuk memastikan kualitas data sebelum digunakan dalam proses klasifikasi. Proses ini meliputi data cleaning, normalisasi, dan encoding label. Data cleaning dilakukan dengan menghapus data duplikat serta menangani missing value sehingga tidak terdapat nilai kosong pada dataset.

Selanjutnya dilakukan normalisasi menggunakan metode Min-Max Scaling untuk mengubah rentang nilai fitur menjadi skala 0 hingga 1. Hal ini bertujuan agar tidak terjadi dominasi nilai pada fitur tertentu. Selain itu, dilakukan encoding label untuk mengubah kelas kualitas TBS dari bentuk kategorikal menjadi numerik, yaitu lewat matang (0), matang (1), dan mentah (2). Contoh hasil preprocessing data dapat dilihat pada Tabel 1.

Tabel 1. Contoh Hasil Preprocessing Data

No	Berat Tandan	Kandungan Minyak	Kadar Air	Brondol Lepas	Kualitas (Label)
1	0.65	0.70	0.45	0.30	1
2	0.40	0.35	0.60	0.20	0
3	0.80	0.75	0.50	0.85	2
4	0.55	0.60	0.48	0.40	1
5	0.35	0.30	0.65	0.15	0

Berdasarkan Tabel 1, terlihat bahwa seluruh nilai fitur telah berada pada rentang 0 hingga 1, yang menunjukkan bahwa proses normalisasi telah berhasil dilakukan. Selain itu, label kualitas telah dikonversi ke dalam bentuk numerik sehingga dapat diproses oleh algoritma SVM.

Hasil Pembagian Data

Setelah tahap preprocessing selesai, data dibagi menjadi dua bagian yaitu data training dan data testing dengan perbandingan 80:20. Dari total 1.500 data, sebanyak 1.200 data digunakan sebagai data training dan 300 data sebagai data testing.

Pembagian data dilakukan menggunakan metode stratified sampling untuk menjaga keseimbangan distribusi kelas pada kedua dataset. Hal ini penting agar model dapat mempelajari pola dari setiap kelas secara proporsional. Distribusi data setelah pembagian dapat dilihat pada Tabel 2.

Tabel 2. Distribusi Data Training dan Testing

No	Kelas	Data Training	Data Testing
1	Mentah	363	105
2	Matang	483	133
3	Lewat Matang	354	62
	Total	1200	300

Berdasarkan Tabel 2, dapat dilihat bahwa distribusi data pada masing-masing kelas tetap seimbang baik pada data training maupun data testing. Hal ini menunjukkan bahwa proses pembagian data telah dilakukan dengan baik sehingga tidak terjadi ketimpangan kelas yang dapat mempengaruhi performa model.

Klasifikasi

Pada tahap ini dilakukan pengujian model Support Vector Machine (SVM) menggunakan data testing sebanyak 300 data. Tujuan pengujian ini adalah untuk mengetahui kemampuan model dalam mengklasifikasikan kualitas Tandan Buah Segar (TBS) ke dalam tiga kelas, yaitu Lewat Matang, Matang, dan Mentah. Hasil klasifikasi disajikan dalam bentuk classification report yang memuat nilai precision, recall, F1-score, dan support untuk masing-masing kelas. Nilai-nilai tersebut digunakan untuk mengevaluasi performa model secara lebih rinci.

Tabel 3. Hasil Klasifikasi (Classification Report)

No	Kelas	Precision	Recall	F1-Score	Support
1	Lewat Matang	0.97	0.93	0.95	105
2	Matang	0.94	0.99	0.97	133
3	Mentah	0.93	0.89	0.91	62

Berdasarkan Tabel 3, performa model dapat dianalisis pada masing-masing kelas. Pada kelas Lewat Matang, model memiliki nilai precision sebesar 0.97 yang menunjukkan bahwa sebagian besar data yang diprediksi sebagai Lewat Matang adalah benar. Namun, nilai recall sebesar 0.93 mengindikasikan bahwa masih terdapat beberapa data Lewat Matang yang tidak terdeteksi dengan baik oleh model, yang kemungkinan disebabkan oleh kemiripan karakteristik dengan kelas Matang. Pada kelas Matang, model menunjukkan performa terbaik dengan nilai recall sebesar 0.99, yang berarti hampir seluruh data pada kelas ini berhasil dikenali dengan benar. Nilai precision sebesar 0.94 menunjukkan bahwa terdapat sedikit data dari kelas lain yang diprediksi sebagai Matang, namun secara keseluruhan F1-score sebesar 0.97 menandakan performa yang sangat baik. Sementara itu, pada kelas Mentah, model memperoleh nilai precision sebesar 0.93 dan recall sebesar 0.89, yang merupakan nilai recall terendah dibandingkan kelas lainnya. Hal ini menunjukkan bahwa model masih mengalami kesulitan dalam mengenali seluruh data pada kelas Mentah, yang kemungkinan disebabkan oleh jumlah data yang lebih sedikit serta adanya kemiripan fitur dengan kelas lainnya.

Evaluasi Model

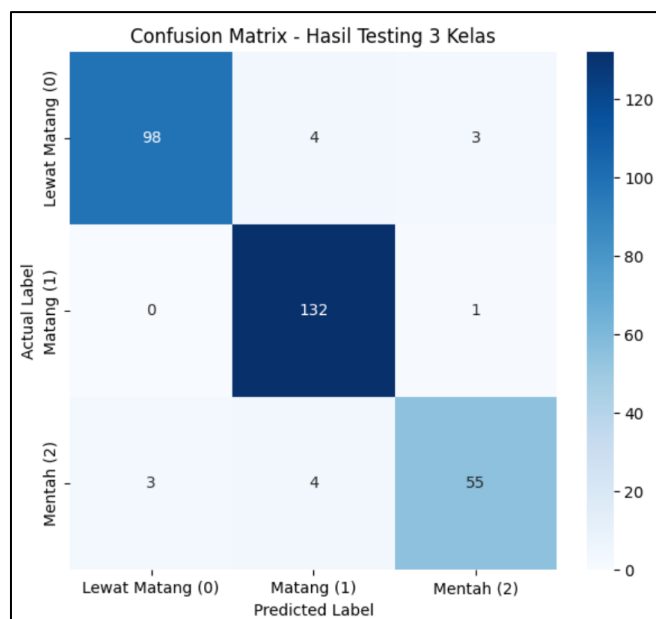
Selain analisis per kelas, dilakukan evaluasi secara keseluruhan untuk mengetahui performa model menggunakan metrik akurasi, macro average, dan weighted average.

Tabel 4. Hasil Evaluasi Model

No	Metode Evaluasi	Nilai
1	Accuracy	0.95
2	Macro Avg Precision	0.95

3	Macro Avg Recall	0.94
4	Macro Avg F1-Score	0.94
5	Weighted Avg Precision	0.95
6	Weighted Avg Recall	0.95
7	Weighted Avg F1-Score	0.95

Berdasarkan Tabel 4, model menghasilkan nilai akurasi sebesar 0.95 (95%), yang menunjukkan bahwa sebagian besar data berhasil diklasifikasikan dengan benar. Nilai macro average menunjukkan rata-rata performa model pada setiap kelas tanpa mempertimbangkan jumlah data, di mana hasil yang mendekati 1 mengindikasikan bahwa model memiliki performa yang seimbang pada semua kelas. Sementara itu, nilai weighted average mempertimbangkan jumlah data pada setiap kelas, sehingga memberikan gambaran performa model secara keseluruhan dengan mempertimbangkan proporsi data. Nilai weighted average yang tetap tinggi menunjukkan bahwa model tidak mengalami bias terhadap kelas tertentu meskipun terdapat perbedaan jumlah data antar kelas.



Gambar 3. Grafik Confusion Matrix

Confusion matrix yang ditampilkan dalam Gambar 3. menunjukkan hasil klasifikasi SVM terhadap tiga kelas kualitas TBS, yaitu Lewat Matang, Matang, dan Mentah. Dari 300 data uji, sebanyak 285 data berhasil diprediksi dengan benar. Secara rinci, 98 dari 105 data aktual Lewat Matang diprediksi benar (4 salah ke Matang, 3 ke Mentah); 132 dari 133 data aktual Matang diprediksi benar (1 salah ke Mentah); dan 55 dari 62 data aktual Mentah diprediksi benar (3 salah ke Lewat Matang, 4 ke Matang). Secara matematis, nilai accuracy mencapai 95 persen, dengan precision, recall, dan F1-score yang tinggi untuk setiap kelas. Hasil pengujian model terhadap data testing ditunjukkan pada confusion matrix berikut:

Tabel 5. Confusion Matrix Hasil Klasifikasi

Aktual	Lewat Matang (0)	Matang (1)	Mentah (2)
Lewat Matang (0)	98	4	3
Matang (1)	0	132	1
Mentah (2)	3	4	55

Berdasarkan Tabel 5, model menunjukkan kemampuan klasifikasi yang sangat baik dengan jumlah prediksi benar yang dominan pada setiap kelas. Kesalahan klasifikasi relatif kecil dan terjadi pada data yang berada di batas antar kelas (boundary region), khususnya antara kelas matang dan lewat matang. Hal ini menunjukkan bahwa karakteristik kedua kelas tersebut memiliki kemiripan fitur yang cukup tinggi.

$$Akurasi = \frac{98 + 132 + 55}{300} = \frac{285}{300} = 0,95$$

a. Kelas Lewat Matang

$$Precision = \frac{98}{98 + 3} = 0,97$$

$$Recall = \frac{98}{98 + 4 + 3} = 0,93$$

$$F1 - score = 2 \times \frac{0,97 \times 0,93}{0,97 + 0,93} = 0,95$$

Pada kelas Lewat Matang, *confusion matrix* menunjukkan bahwa dari 105 data aktual Lewat Matang, sebanyak 98 data diprediksi dengan benar sebagai Lewat Matang, 4 data diprediksi keliru sebagai Matang, dan 3 data diprediksi keliru sebagai Mentah. Artinya, hanya 7 data dari kelas ini yang mengalami kesalahan klasifikasi. Nilai *recall* sebesar 93% mencerminkan kemampuan model mengenali sebagian besar data Lewat Matang. Sementara itu, dari sisi kolom prediksi Lewat Matang, tidak ada data dari kelas Matang yang salah masuk ke kelas ini, dan hanya 3 data aktual Mentah yang keliru diprediksi sebagai Lewat Matang, sehingga *precision* mencapai 97%.

b. Kelas Matang

$$Precision = \frac{132}{132 + 8} = 0,94$$

$$Recall = \frac{132}{132 + 1} = 0,99$$

$$F1 - score = 2 \times \frac{0,94 \times 0,99}{0,94 + 0,99} = 0,97$$

Pada kelas Matang, *confusion matrix* menunjukkan bahwa dari 133 data aktual Matang, sebanyak 132 data diprediksi dengan benar sebagai Matang dan hanya 1 data yang diprediksi keliru sebagai Mentah. Ini merupakan kelas dengan kesalahan terkecil, hanya 1 data yang salah. Nilai *recall* mencapai 99%, yang berarti hampir seluruh data Matang berhasil dikenali dengan tepat. Namun dari prediksi Matang, ada 4 data aktual Lewat Matang dan 4 data aktual Mentah yang ikut diprediksi sebagai Matang, sehingga total prediksi Matang menjadi 140 dan *precision* sebesar 94%.

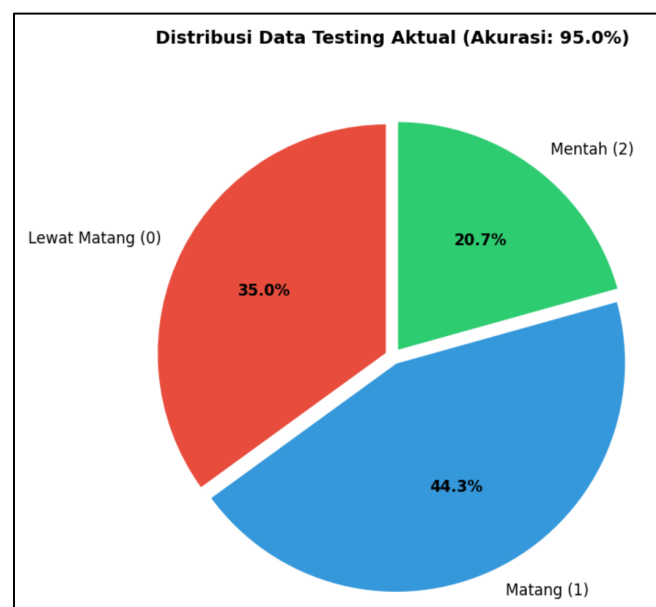
c. Kelas Mentah

$$Precision = \frac{55}{55 + 4} = 0,93$$

$$Recall = \frac{55}{55 + 3 + 4} = 0,89$$

$$F1 - score = 2 \times \frac{0,93 \times 0,89}{0,93 + 0,89} = 0,91$$

Pada kelas Mentah, *confusion matrix* menunjukkan bahwa dari 62 data aktual Mentah, sebanyak 55 data diprediksi dengan benar sebagai Mentah, 3 data diprediksi keliru sebagai Lewat Matang, dan 4 data diprediksi keliru sebagai Matang. Total 7 data mengalami kesalahan klasifikasi, yang menjadikan kelas ini memiliki tingkat misklasifikasi tertinggi dibanding dua kelas lainnya. Nilai *recall* sebesar 89% mencerminkan adanya tumpang tindih karakteristik fitur antara buah Mentah dengan kedua kelas lainnya. Dari sisi kolom prediksi Mentah, hanya 3 data aktual Lewat Matang dan 1 data aktual Matang yang keliru masuk ke kelas ini, sehingga *precision* mencapai 93%.



Gambar 4. Distribusi data

Gambar 4. menampilkan distribusi data uji (*testing data*) aktual sebanyak 300 sampel yang digunakan dalam proses evaluasi model. Berdasarkan diagram tersebut, terlihat bahwa kelas Matang mendominasi dengan proporsi terbesar yaitu 44,3% atau sebanyak 133 data, diikuti oleh kelas Lewat Matang sebesar 35,0% atau 105 data, dan kelas Mentah sebesar 20,7% atau 62 data. Distribusi ini mencerminkan komposisi alami kualitas TBS di lapangan, di mana buah yang berada pada fase matang optimal cenderung lebih dominan dalam hasil panen harian PT XYZ periode Februari 2025.

4. Kesimpulan

Berdasarkan hasil penelitian dan pembahasan yang telah dilakukan, dapat disimpulkan bahwa metode Support Vector Machine (SVM) dengan kernel Radial Basis Function (RBF) mampu mengklasifikasikan kualitas Tandan Buah Segar (TBS) kelapa sawit secara efektif ke dalam tiga kategori, yaitu Lewat Matang, Matang, dan Mentah. Model yang dibangun menggunakan 1.500 dataset dengan pembagian 80:20 menghasilkan akurasi sebesar 95%, di mana 285 dari 300 data uji berhasil diklasifikasikan dengan benar. Hasil evaluasi menunjukkan bahwa kelas Matang memiliki performa terbaik dengan nilai recall sebesar 99%, diikuti oleh kelas Lewat Matang dengan precision 97% dan recall 93%, sedangkan kelas Mentah memiliki recall terendah sebesar 89% akibat adanya kemiripan karakteristik fitur antar kelas. Proses preprocessing menggunakan Min-Max Scaling terbukti mampu meningkatkan performa model dengan menyeimbangkan skala data. Hasil confusion matrix menunjukkan bahwa kesalahan klasifikasi relatif kecil dan umumnya terjadi pada data yang berada di batas antar kelas (boundary region), khususnya antara kelas Matang dan Lewat Matang. Secara keseluruhan, model SVM yang dihasilkan memiliki performa yang baik, stabil, dan tidak bias terhadap distribusi kelas, sehingga dapat digunakan sebagai solusi untuk meningkatkan objektivitas dan efisiensi dalam proses klasifikasi kualitas TBS di industri kelapa sawit. Oleh karena itu, penelitian selanjutnya disarankan untuk mengembangkan model ini ke dalam aplikasi berbasis mobile atau Internet of Things (IoT) agar dapat digunakan langsung di lapangan, melakukan optimasi hyperparameter seperti nilai C dan gamma untuk meningkatkan performa model, serta menambahkan fitur lain seperti analisis citra digital guna meningkatkan akurasi klasifikasi secara lebih komprehensif.

Referensi

- [1] T. Gori *et al.*, "PREPROCESSING DATA DAN KLASIFIKASI UNTUK PREDIKSI KINERJA DATA PREPROCESSING AND CLASSIFICATION FOR PREDICTING STUDENT," vol. 11, no. 1, pp. 215–224, 2024, doi: 10.25126/jtiik.20241118074.
- [2] E. Hokijuliandy, H. Napitupulu, and F. Firdaniza, "Analisis Sentimen Menggunakan Metode Klasifikasi Support Vector Machine (SVM)," vol. 5, no. 2, pp. 40–49, 2023, doi: 10.37278/sisinfo.v5i2.670.
- [3] P. P. Allorerung, A. Erna, and M. Bagussahri, "Analisis Performa Normalisasi Data untuk Klasifikasi K-Nearest Neighbor pada Dataset Penyakit," vol. 9, no. 3, pp. 178–191, 2024, doi: 10.14421/jiska.2024.9.3.
- [4] H. S. Wafa *et al.*, "Prediksi Penyakit Diabetes Menggunakan Algoritma Support Vector Machine (SVM)," vol. 1, pp. 40–45, 2022, doi: 10.36423/index.v4i1.895.
- [5] A. Huesesan and M. A. Khoiri, "Kajian Komponen Produksi Tanaman Kelapa Sawit (Elaeis guineensis Jacq.) yang diberi Pupuk Boron dan NPK," vol. 36, no. 1, pp. 168–181, 2025, doi: 10.24198/agrikultura.v36i1.55965.
- [6] B. N. Azmi, A. Hermawan, and D. Avianto, "Analisis Pengaruh Komposisi Data Training dan Data Testing pada Penggunaan PCA dan Algoritma Decision Tree untuk Klasifikasi Penderita Penyakit Liver," vol. 4, no. 4, pp. 281–290, 2023, doi: 10.35746/jtim.v4i4.298.
- [7] A. Jalil, A. Homaidi, and Z. Fatah, "Implementasi Algoritma Support Vector Machine Untuk Klasifikasi Status Stunting Pada Balita," vol. 8, no. 3, pp. 2070–2079, 2024, doi: 10.33379/gtech.v8i3.4811.
- [8] M. Basyuni, N. Amri, L. Agustina, and P. Putri, "Characteristics of Fresh Fruit Bunch Yield and the Physicochemical Qualities of Palm Oil during Storage in North Sumatra, Indonesia," vol. 17, no. 2, pp. 182–190, 2022, doi: 10.22146/ijc.20210.
- [9] M. A. Yolov, R. Kurniawan, A. T. Martadinata, and S. D. Cahyo, "Klasifikasi Tingkat Kematangan Buah Sawit Berbasis Deep Learning dengan Klasifikasi Tingkat Kematangan Buah Sawit Berbasis Deep Learning dengan Menggunakan Arsitektur Yolov5," no. October 2023, 2024, doi: 10.47065/josh.v5i1.4408.
- [10] N. Sari *et al.*, "Klasifikasi Tingkat Kematangan Tandan Buah Segar Kelapa Sawit Menggunakan Probe Optik Ripeness Classification of Oil Palm Fresh Fruit Bunch Using an Optical Probe," vol. 8, no. 3, pp. 72–77, 2022, doi: 10.24815/jacps.v8i3.14122.
- [11] S. C. Bilqisth and R. M. Ikhsanuddin, "ANALISIS PERBANDINGAN AKURASI KLASIFIKASI KEPUASAN SISWA TERHADAP KINERJA GURU," vol. 20, pp. 37–44, 2025, doi: 10.30587/e-link.v20i1.9719.
- [12] M. R. Kusnaldi, T. Gulo, and S. Aripin, "Penerapan Normalisasi Data Dalam Mengelompokkan Data Mahasiswa Dengan Menggunakan Metode K-Means Untuk Menentukan Prioritas Bantuan Uang Kuliah Tunggal," vol. 3, no. 4, pp. 330–338, 2022, doi: 10.47065/josyc.v3i4.2112.
- [13] S. L. Abdullah and A. F. Akbariyah, "POTENSI EKSPOR CRUDE PALM OIL (CPO) DI INDONESIA," vol. 4307, no. 1, pp. 61–67, 2024, doi: 10.54314/jssr.v7i1.1713.
- [14] L. N. Farida and S. Bahri, "Klasifikasi Gagal Jantung Menggunakan Metode SVM (Support Vector Machine)," *Komputika J. Sist. Komput.*, vol. 13, no. 2, pp. 149–156, 2024, doi: 10.34010/komputika.v13i2.11330.
- [15] R. S. Nurhalizah and R. Ardianto, "Analisis Supervised dan Unsupervised Learning pada Machine Learning: Systematic Literature Review," vol. 4, no. 1, pp. 61–72, 2024, doi: 10.54082/jiki.168.
- [16] Ramadhan, Andarsyah, 2022, KLASIFIKASI TEXT SPAM MENGGUNAKAN METODE SUPPORT VECTOR MACHINE DAN NAÏVE BAYES."