

Sistem Rekomendasi Jenis Perangsang Buah Mangga Berdasarkan Data Fisiologis Tanaman Menggunakan Algoritma Naive Bayes

^{1*}Pendi Maulana, ²Daniel Swanjaya, ³Resty Wulanningrum

¹²³Teknik Informatika, Universitas Nusantara PGRI Kediri

E-mail: ¹pendimaaulana001@gmail.com, ²daniel@unpkediri.ac.id, ³restyw@unpkdr.ac.id

Penulis Korespondensi : Pendi Maulana

Abstrak—Peningkatan produktivitas tanaman mangga dapat dilakukan melalui pemberian perangsang buah yang tepat. Namun, dalam praktiknya, penentuan jenis dan dosis perangsang masih sering dilakukan berdasarkan pengalaman sehingga hasilnya belum optimal. Penelitian ini bertujuan untuk mengembangkan sistem rekomendasi berbasis data dengan memanfaatkan metode *machine learning*. Data yang digunakan mencakup kondisi *fisiologis* tanaman, seperti kondisi daun, *fase* pertumbuhan, tingkat serangan hama, serta tingkat kerontokan buah. Metode yang digunakan adalah *klasifikasi* menggunakan algoritma *Naive Bayes*, *Logistic Regression*, dan *Gaussian Mixture Model*. Dataset terdiri dari 350 data asli dan 1100 data sintetis untuk meningkatkan keseimbangan data. Hasil penelitian menunjukkan bahwa algoritma *Naive Bayes* memiliki performa terbaik dengan peningkatan akurasi dari 66,43% menjadi 76,59% setelah penggunaan data sintetis. Sistem rekomendasi yang dihasilkan diharapkan dapat membantu petani mangga dalam menentukan jenis dan dosis perangsang buah secara lebih tepat dan berbasis data sehingga meningkatkan produktivitas secara optimal.

Kata Kunci— data fisiologis, klasifikasi, *machine learning*, mangga, rekomendasi

Abstract— *Mango plant productivity can be increased by administering appropriate fruit stimulants. However, in practice, the determination of the type and dosage of stimulants is often based on experience, resulting in suboptimal results. This study aims to develop a data-driven recommendation system utilizing machine learning methods. The data used includes plant physiological conditions, such as leaf condition, growth phase, pest attack level, and fruit loss rate. The methods used are classification using the Naive Bayes algorithm, Logistic Regression, and Gaussian Mixture Model. The dataset consists of 350 original data and 1100 synthetic data to improve data balance. The results showed that the Naive Bayes algorithm had the best performance with an accuracy increase from 66.43% to 76.59% after the use of synthetic data. The resulting recommendation system is expected to assist mango farmers in determining the type and dosage of fruit stimulants more precisely and based on data, thereby optimally increasing productivity.*

Keywords— classification, machine learning, mango, physiological data, recommendation

This is an open access article under the CC BY-SA License.



I. PENDAHULUAN

Tanaman mangga merupakan salah satu komoditas hortikultura yang memiliki nilai ekonomi tinggi dan banyak dibudidayakan di Indonesia[1]. Tingginya permintaan pasar terhadap buah mangga mendorong petani untuk meningkatkan produktivitas hasil panen[2]. Salah satu upaya yang dilakukan adalah penggunaan perangsang buah untuk mempercepat

proses pembungaan dan pembuahan[3]. Namun, keberhasilan penggunaan perangsang tersebut sangat bergantung pada kondisi fisiologis tanaman[4].

Dalam praktiknya, penentuan jenis dan dosis perangsang buah masih sering dilakukan secara subjektif berdasarkan pengalaman petani. Hal ini menyebabkan hasil yang diperoleh kurang optimal dan tidak konsisten. Permasalahan ini dapat diatasi dengan pendekatan berbasis data, yaitu dengan memanfaatkan *machine learning* untuk mengolah data kondisi tanaman sehingga menghasilkan rekomendasi yang lebih objektif dan akurat[5].

Penelitian sebelumnya menunjukkan bahwa *machine learning* dapat meningkatkan kualitas pengambilan keputusan berbasis data. menyatakan bahwa pemanfaatan kecerdasan buatan mampu meningkatkan efisiensi analisis data[6]. Selain itu, menunjukkan bahwa pendekatan berbasis data dapat menghasilkan prediksi yang lebih akurat dalam bidang pertanian[7].

Berdasarkan hal tersebut, penelitian ini bertujuan untuk mengembangkan model klasifikasi yang mampu memberikan rekomendasi perangsang buah berdasarkan kondisi fisiologis tanaman mangga. Berbeda dengan penelitian sebelumnya yang umumnya hanya menggunakan satu pendekatan klasifikasi terbimbing, penelitian ini membandingkan model berbasis *probabilitas* terkaster seperti *Gaussian Mixture Model* dengan model klasifikasi terbimbing seperti *Naive Bayes* dan *Logistic Regression*. Perbandingan ini penting karena data *fisiologis* tanaman mangga memiliki karakteristik kompleks, tidak selalu terpisah secara *linier*, serta berpotensi membentuk pola distribusi tersembunyi. Oleh karena itu, pendekatan berbasis pengklasteran *probabilistik* dapat menangkap struktur data laten, sementara model klasifikasi terbimbing memberikan kemampuan prediksi berbasis label. Analisis komparatif ini diharapkan dapat mengidentifikasi metode yang paling sesuai untuk karakteristik data *fisiologis* tanaman mangga. Diharapkan hasil penelitian ini dapat membantu petani dalam mengambil keputusan yang lebih tepat dan meningkatkan produktivitas tanaman.

II. METODE

Penelitian ini menggunakan pendekatan eksperimen dengan tujuan untuk membangun dan mengevaluasi model klasifikasi dalam memberikan rekomendasi perangsang buah pada tanaman mangga[8]. Data yang digunakan dalam penelitian ini merupakan data kondisi *fisiologis* tanaman yang diperoleh melalui observasi lapangan. Parameter yang diamati meliputi kondisi daun, *fase* perkembangan tanaman, tingkat serangan hama, serta tingkat kerontokan buah.

Dataset yang digunakan terdiri dari dua jenis, yaitu data asli sebanyak 350 data dan data sintetis sebanyak 1100 data. Data sintetis dibangkitkan menggunakan pendekatan *rule-based probabilistic augmentation* untuk meningkatkan jumlah data serta memperbaiki distribusi kelas agar lebih seimbang.

Secara algoritmik, proses pembangkitan data sintetis dilakukan melalui beberapa tahapan. Pertama, dilakukan ekstraksi aturan (*rule extraction*) dari data asli menggunakan pendekatan *modus*, yaitu:

1. Untuk setiap atribut target Y dan fitur X, dibentuk aturan:

$$f(X) = \text{model}(Y | X) \quad (1)$$

Sebagai contoh:

- 1) aturan obat berdasarkan serangan hama

- 2) aturan dosis berdasarkan tingkat serangan
- 3) aturan cara pakai berdasarkan kondisi daun

Kedua, dilakukan proses *sampling* dengan mengambil satu data acak dari dataset asli, kemudian dimodifikasi menggunakan aturan yang telah diperoleh. Penerapan aturan dilakukan secara probabilistik dengan peluang sebesar $p=0.7$, sehingga:

$$X_{baru} = \{f(X), X_{asli}, \text{dengan probabilitas } 0.7 \text{ dengan probabilitas } 0.3\} \quad (2)$$

Ketiga, proses ini diulang hingga jumlah data sintetis yang dihasilkan mencapai target yang diinginkan. Data sintetis kemudian digabungkan dengan data asli dan dilakukan pengacakan (*shuffling*) untuk menghindari bias urutan data.

Tahapan penelitian dimulai dari pengumpulan data, *preprocessing* data, pembagian dataset, pelatihan model, hingga evaluasi model. Proses *preprocessing* meliputi pembersihan data, transformasi data kategorikal menjadi numerik, serta normalisasi data jika diperlukan. Dataset kemudian dibagi menjadi data latih sebesar 80% dan data uji sebesar 20%.

Metode klasifikasi yang digunakan dalam penelitian ini adalah *Naive Bayes*, *Logistic Regression*, dan *Gaussian Mixture Model*. Evaluasi model dilakukan menggunakan *confusion matrix* dengan metrik *akurasi*, *presisi*, *recall*, dan *F1-score*[9].

A. Rumus Matematika

1. Naive Bayes

$$P(C|X) = \frac{P(X|C) \cdot P(C)}{P(X)} \quad (3)$$

2. Logistic Regression

$$P(y = 1|x) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x)}} \quad (4)$$

3. Gaussian Mixture Model

$$p(x) = \sum_{k=1}^K \pi_k \cdot \mathcal{N}(x|\mu_k, \Sigma_k) \quad (5)$$

4. Accuracy

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (6)$$

5. Precision

$$Precision = \frac{TP}{TP + FP} \quad (7)$$

6. Recall

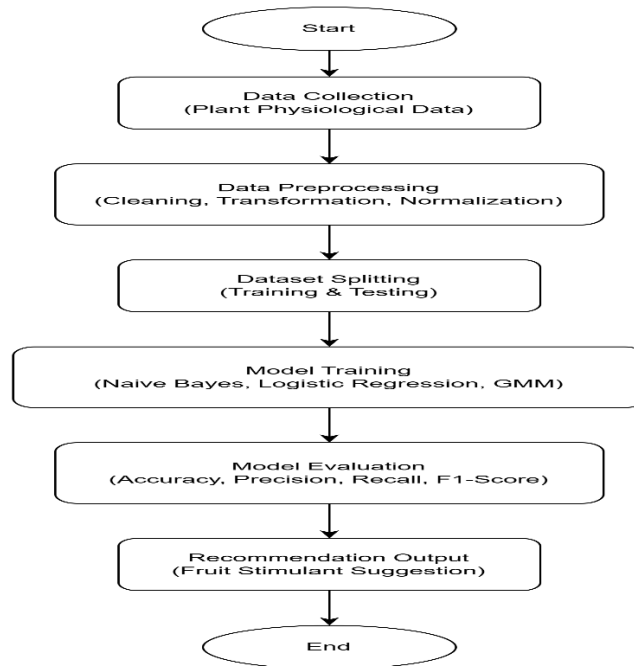
$$Recall = \frac{TP}{TP + FN} \quad (8)$$

7. F1-Score

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \quad (9)$$

Keterangan:

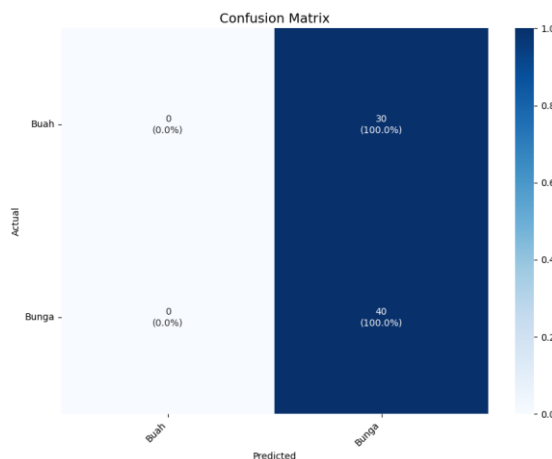
TP (*True Positive*), TN (*True Negative*), FP (*False Positive*), dan FN (*False Negative*) merupakan komponen dalam *confusion matrix* yang digunakan untuk mengevaluasi performa model.



Gambar 1. Diagram Alur Penelitian

Gambar 1 menunjukkan alur penelitian yang dimulai dari proses pengumpulan data, preprocessing, pembagian dataset, pelatihan model, hingga evaluasi model untuk menghasilkan rekomendasi perangsang buah.

III. HASIL DAN PEMBAHASAN



Gambar 2. *Gaussian Mixture Model* (Data Asli)

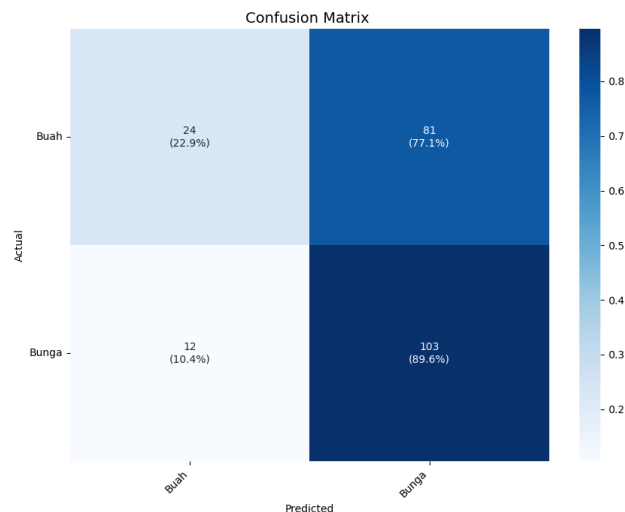
Berdasarkan Gambar 2 visualisasi confusion matrix, terlihat bahwa model cenderung mengklasifikasikan seluruh data ke dalam satu kelas, yaitu kelas bunga. Hal ini ditunjukkan oleh dominasi nilai pada kolom prediksi bunga, sementara kolom prediksi buah hampir tidak memiliki nilai.

Kondisi tersebut menunjukkan bahwa model belum mampu membedakan karakteristik antar kelas dengan baik. Akibatnya, baik data yang sebenarnya termasuk dalam kelas buah

maupun bunga diprediksi sebagai bunga. Hal ini berdampak pada rendahnya nilai *akurasi*, *precision*, *recall*, dan *F1-score* yang dihasilkan.

Fenomena ini disebabkan oleh distribusi data yang tidak seimbang, sehingga model kesulitan mempelajari pola dari kelas yang jumlah datanya lebih sedikit. Selain itu, sifat *Gaussian Mixture Model* sebagai metode berbasis kluster (*unsupervised*) juga menyebabkan proses pemetaan klaster ke label kelas menjadi kurang akurat.

Secara keseluruhan, hasil ini menunjukkan bahwa model belum mampu menangkap pola data secara optimal pada kondisi data asli. Oleh karena itu, diperlukan perbaikan kualitas data, seperti penyeimbangan distribusi atau penambahan data sintesis.



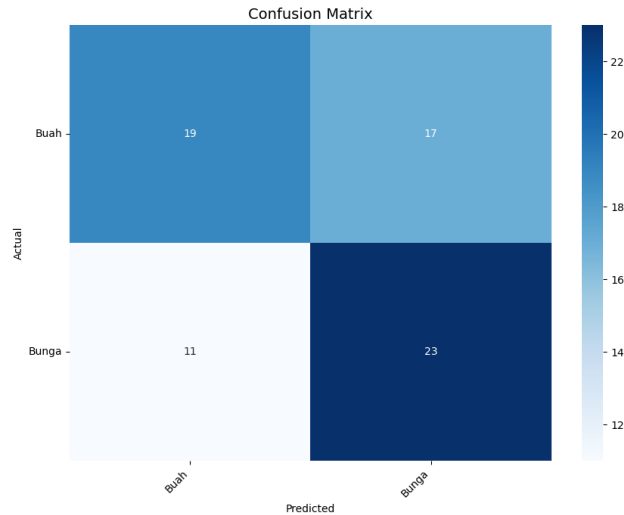
Gambar 3. *Gaussian Mixture Model* (Data Sintesis)

Gambar 3 menunjukkan Setelah menggunakan data sintesis, *confusion matrix* menunjukkan adanya peningkatan kemampuan model dalam mengenali kedua kelas. Model mulai mampu mengklasifikasikan sebagian data buah dan bunga dengan benar, meskipun masih terdapat kesalahan prediksi yang cukup signifikan.

Dibandingkan dengan kondisi sebelumnya, model tidak lagi mengelompokkan seluruh data ke dalam satu kelas. Hal ini menunjukkan bahwa penambahan data sintesis membantu model dalam mengenali pola distribusi data yang lebih beragam.

Namun demikian, tingkat kesalahan klasifikasi masih tergolong tinggi, khususnya pada kelas buah yang masih sering diprediksi sebagai bunga. Hal ini menunjukkan bahwa meskipun distribusi data telah membaik, model masih memiliki keterbatasan dalam menangkap pola klasifikasi secara akurat.

Secara umum, penggunaan data sintesis memberikan dampak positif terhadap stabilitas dan performa model, meskipun peningkatannya belum signifikan. Oleh karena itu, diperlukan pendekatan model yang lebih sesuai untuk meningkatkan hasil klasifikasi.



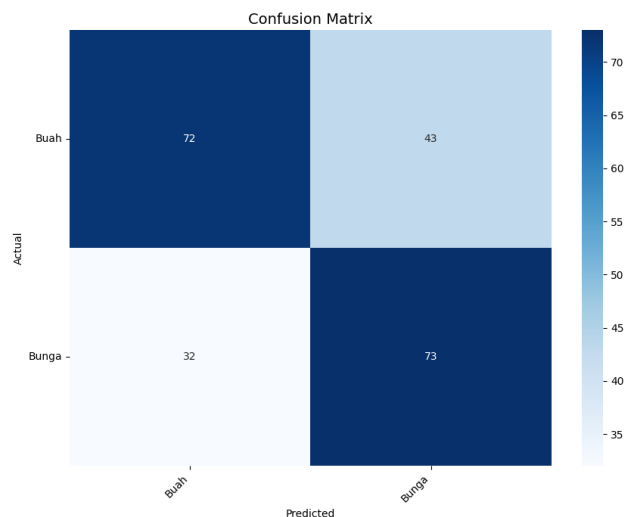
Gambar 4. *Logistic Regression* (Data Asli)

Berdasarkan Gambar 4 *confusion matrix*, model *Logistic Regression* telah mampu mengklasifikasikan kedua kelas dengan cukup baik. Sebagian besar data berhasil diprediksi sesuai dengan kelas sebenarnya, meskipun masih terdapat beberapa kesalahan klasifikasi.

Distribusi prediksi yang dihasilkan terlihat cukup seimbang antara kedua kelas, yang menunjukkan bahwa model tidak memiliki kecenderungan dominan terhadap salah satu kelas tertentu. Hal ini mencerminkan kestabilan model dalam melakukan prediksi.

Keseimbangan ini juga tercermin pada nilai *akurasi*, *precision*, *recall*, dan *F1-score* yang relatif berdekatan. Dengan demikian, model tidak hanya memiliki tingkat akurasi yang cukup baik, tetapi juga mampu mengenali masing-masing kelas secara proporsional.

Meskipun demikian, masih terdapat kesalahan klasifikasi yang menunjukkan bahwa model belum sepenuhnya mampu menangkap seluruh pola dalam data, terutama pada data dengan karakteristik yang mirip antar kelas.

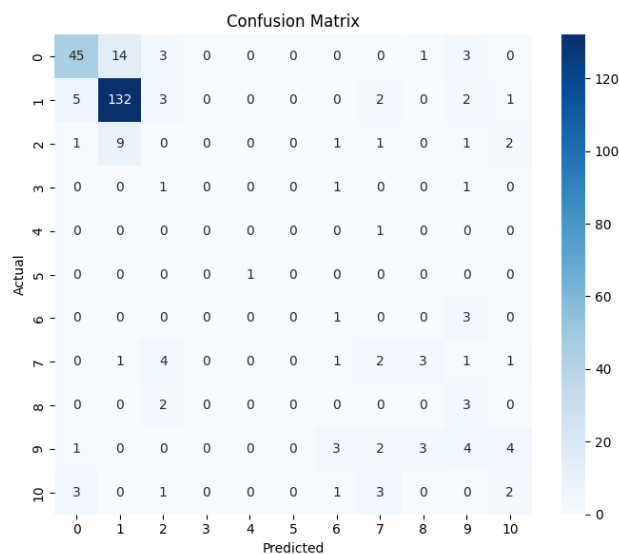


Gambar 5. *Logistic Regression* (Data Sintetis)

Gambar 5 menunjukkan Setelah dilakukan penambahan data sintetis, *confusion matrix* menunjukkan peningkatan kemampuan model dalam mengklasifikasikan kedua kelas. Jumlah

prediksi yang benar pada masing-masing kelas mengalami peningkatan dibandingkan dengan kondisi sebelumnya.

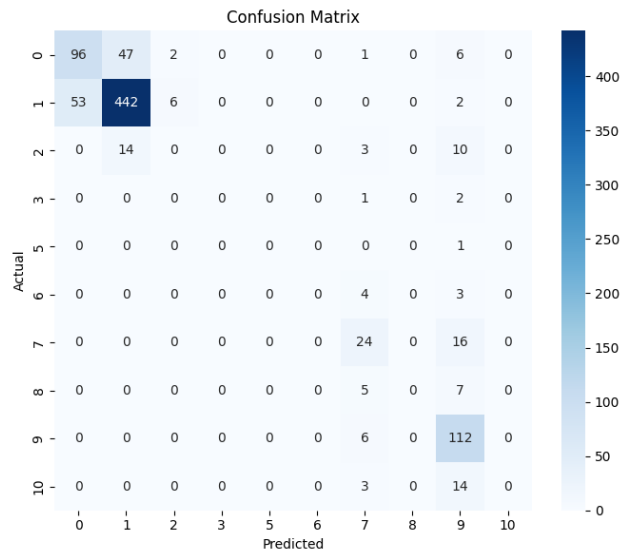
Distribusi prediksi menjadi lebih merata dan konsisten, yang menunjukkan bahwa model mampu memanfaatkan data sintetis untuk mengenali pola yang sebelumnya kurang terlihat pada data asli. Peningkatan ini juga terlihat pada metrik evaluasi, di mana nilai *akurasi*, *precision*, *recall*, dan *F1-score* mengalami kenaikan. Hal ini menunjukkan bahwa model tidak hanya menjadi lebih akurat, tetapi juga lebih seimbang dalam mengenali setiap kelas. Walaupun masih terdapat beberapa kesalahan prediksi, secara keseluruhan performa model mengalami peningkatan yang cukup signifikan setelah penggunaan data sintetis.



Gambar 6. *Naive Bayes* (Data Asli)

Gambar 6 *Confusion matrix* menunjukkan bahwa sebagian besar nilai berada pada diagonal utama, yang menandakan bahwa model mampu menghasilkan prediksi yang benar dalam jumlah yang cukup besar untuk setiap kelas. Nilai di luar diagonal relatif kecil, yang berarti tingkat kesalahan klasifikasi tidak terlalu signifikan. Hal ini menjadi salah satu faktor yang menyebabkan nilai *akurasi*, *recall*, dan *F1-score* model ini lebih tinggi dibandingkan model lainnya pada data asli.

Selain itu, keseimbangan antar metrik evaluasi menunjukkan bahwa model tidak hanya unggul dalam satu aspek, tetapi mampu menjaga konsistensi performa secara keseluruhan. Dengan demikian, dapat disimpulkan bahwa *Naive Bayes* memiliki kemampuan yang baik dalam memahami pola data meskipun jumlah data masih terbatas.



Gambar 7. *Naive Bayes* (Data Sintetis)

Gambar 7 menunjukkan Setelah menggunakan data sintetis, terlihat bahwa nilai pada diagonal utama semakin dominan. Hal ini menunjukkan bahwa jumlah prediksi yang benar meningkat secara signifikan pada hampir seluruh kelas. Kesalahan prediksi yang ditunjukkan oleh nilai di luar diagonal menjadi lebih terkontrol dan tidak menyebar luas. Hal ini menunjukkan bahwa model menjadi lebih konsisten dalam melakukan klasifikasi.

Peningkatan ini menunjukkan bahwa kualitas dan distribusi data yang lebih baik memberikan dampak langsung terhadap performa model. *Naive Bayes* mampu memanfaatkan data tambahan untuk mengenali pola yang sebelumnya belum terlihat dengan jelas.

Dibandingkan dengan model lain, peningkatan performa pada *Naive Bayes* merupakan yang paling signifikan. Hal ini sejalan dengan peningkatan nilai akurasi, *precision*, *recall*, dan *F1-score* yang cukup tinggi. Secara keseluruhan, dapat disimpulkan bahwa penggunaan data sintetis memberikan dampak yang sangat positif terhadap performa model *Naive Bayes*, sehingga menjadikannya sebagai model terbaik dalam penelitian ini.

Tabel 1. Distribusi Dataset

No	Jenis Data	Jumlah
1	Data Asli	350
2	Data Sintesis	1100
	Total	1450

Tabel 1 menunjukkan distribusi dataset yang digunakan dalam penelitian ini, yang terdiri dari data asli dan data sintesis. Data sintesis digunakan untuk menambah jumlah data serta memperbaiki distribusi kelas agar lebih seimbang. Dengan total 1450 data, diharapkan model dapat belajar dengan lebih optimal dan menghasilkan performa yang lebih baik.

Tabel 2. Hasil Evaluasi Model (Data Asli)

No	Model	Accuracy	Precision	Recall	F1-Score
1	Naive Bayes	0.6643	0.6619	0.6643	0.6607

2	Logistic Regression	0.6000	0.6042	0.6021	0.5987
3	Gaussian Mixture Model	0.4429	0.4419	0.4420	0.4418

tabel 2 menunjukkan hasil evaluasi model menggunakan data asli. Terlihat bahwa algoritma *Naive Bayes* memiliki performa terbaik dibandingkan metode lainnya. Namun, secara keseluruhan nilai akurasi masih tergolong sedang. Hal ini disebabkan oleh jumlah data yang terbatas serta distribusi data yang belum seimbang, sehingga memengaruhi kemampuan model dalam mengenali pola secara optimal.

Tabel 3. Hasil Evaluasi Model (Data Sintesis)

No	Model	Accuracy	Precision	Recall	F1-Score
1	Naive Bayes	0.7659	0.7236	0.7659	0.7407
2	Logistic Regression	0.6591	0.6608	0.6607	0.6591
3	Gaussian Mixture Model	0.4909	0.4894	0.4894	0.4894

Tabel 3 menunjukkan hasil evaluasi model setelah dilakukan penambahan data sintesis. Terjadi peningkatan performa pada seluruh model yang digunakan. Hal ini menunjukkan bahwa data sintesis berhasil membantu memperbaiki distribusi data sehingga model dapat belajar dengan lebih baik. *Naive Bayes* tetap menjadi model dengan performa terbaik.

Tabel 4. Perbandingan sebelum vs sesudah (ringkasan)

Model	Akurasi Awal	Akurasi Setelah	Peningkatan
Naive Bayes	0.6643	0.7659	+0.1016
Logistic Regression	0.6000	0.6591	+0.0591
GMM	0.4429	0.4909	+0.0480

Tabel 4 menyajikan perbandingan performa model sebelum dan sesudah penggunaan data sintesis. Terlihat bahwa seluruh model mengalami peningkatan akurasi, dengan peningkatan terbesar terjadi pada algoritma *Naive Bayes*. Hal ini menunjukkan bahwa penambahan data sintesis memberikan dampak signifikan terhadap peningkatan kinerja model.

Hasil penelitian ini diperoleh melalui proses pelatihan dan pengujian model klasifikasi menggunakan tiga algoritma, yaitu *Naive Bayes*, *Logistic Regression*, dan *Gaussian Mixture Model*[10]. Data yang digunakan terdiri dari data asli dan data sintetis yang bertujuan untuk meningkatkan kualitas serta distribusi dataset. Tabel 2 dan Tabel 3 menyajikan nilai *classification report* secara lengkap yang meliputi *akurasi*, *precision*, *recall*, dan *F1-score* untuk masing-masing algoritma sebelum dan sesudah proses *augmentasi* data. Berdasarkan Tabel 2, diketahui bahwa algoritma *Naive Bayes* memiliki performa terbaik dibandingkan metode lainnya pada data asli, sedangkan *Logistic Regression* dan *Gaussian Mixture Model* menunjukkan performa yang lebih rendah. Setelah dilakukan penambahan data sintetis, hasil pada Tabel 3 menunjukkan adanya peningkatan pada seluruh metrik evaluasi, baik *akurasi*, *precision*, *recall*, maupun *F1-score*. Perbandingan peningkatan performa tersebut juga dirangkum pada Tabel 4, yang menunjukkan bahwa seluruh model mengalami peningkatan kinerja setelah *augmentasi* data, dengan peningkatan paling signifikan terjadi pada algoritma *Naive Bayes*.

Selain itu, analisis lebih mendalam terhadap hasil *klasifikasi* dilakukan menggunakan grafik *confusion matrix* yang ditampilkan pada Gambar 2 hingga Gambar 7. Visualisasi tersebut digunakan untuk membedah tingkat akurasi prediksi pada masing-masing kelas dosis atau jenis perangsang buah. Berdasarkan *grafik* tersebut, model dengan performa yang lebih baik menunjukkan dominasi nilai pada diagonal utama, yang menandakan tingginya tingkat prediksi yang sesuai dengan kelas sebenarnya. Sebaliknya, model dengan performa lebih rendah menunjukkan penyebaran nilai di luar diagonal yang lebih besar, yang mengindikasikan masih tingginya kesalahan klasifikasi antar kelas.

Secara umum, hasil penelitian ini menunjukkan bahwa kualitas dan jumlah data sangat berpengaruh terhadap performa model *machine learning*. Penggunaan data sintesis terbukti mampu mengatasi permasalahan ketidakseimbangan data, sehingga model dapat belajar dengan lebih optimal. Hal ini sejalan dengan konsep dasar *machine learning* yang menyatakan bahwa model akan menghasilkan prediksi yang lebih baik jika dilatih menggunakan data yang *representatif* dan seimbang.

IV. KESIMPULAN

Berdasarkan hasil penelitian, penerapan metode machine learning dapat membantu memberikan rekomendasi perangsang buah pada tanaman mangga secara lebih objektif dan tepat, dengan mempertimbangkan kondisi fisiologis tanaman. Dari beberapa model yang diuji, algoritma Naive Bayes menunjukkan kinerja terbaik dibandingkan metode lainnya, dengan peningkatan akurasi dari 66,43% menjadi 76,59% setelah penambahan data sintesis. Hal ini menunjukkan bahwa penggunaan data sintesis mampu memperbaiki distribusi data sehingga meningkatkan kemampuan model dalam melakukan prediksi. Secara keseluruhan, penelitian ini memberikan kontribusi dalam pengembangan sistem rekomendasi berbasis data di bidang pertanian, khususnya untuk membantu petani dalam mengambil keputusan yang lebih efektif dan efisien dalam budidaya tanaman mangga.

DAFTAR PUSTAKA

- [1] E. Ariningsih, N. Ashari, H. P. Saliem, M. Maulana, and K. S. Septanti, "KINERJA AGRIBISNIS MANGGA GEDONG GINCU DAN POTENSINYA SEBAGAI PRODUK EKSPOR PERTANIAN UNGGULAN," *Forum penelitian Agro Ekonomi*, vol. 39, no. 1, p. 49, Sep. 2021, doi: 10.21082/fae.v39n1.2021.49-71.
- [2] A. Sitanggang, D. Raharjo, and A. Muhammad Fajri, "ANALISIS KOMODITAS UNGGULAN DAN STRATEGI PENGEMBANGAN TANAMAN BUAH TAHUNAN DI KABUPATEN CILACAP," *Jurnal Inovasi Daerah*, vol. 4, no. 2, pp. 105–118, Dec. 2025, doi: 10.56655/jid.v4i2.428.
- [3] A. Okta, B. Margita, and W. Tyas, "SIMBIOSIS: Jurnal Sains Pertanian Pengaruh Pemberian POC Tauge Dan Seresah Daun Bagi Pembuahan Jambu Air (*Syzygium aqueum*) Varietas Deli Hijau Dengan Teknik Tabulampot The Effect of Providing POC Tauge and Leaf Litter on the Fruiting of Water Guava (*Syzygium aqueum*) Deli Hijau Variety Using the Tabulampot Technique," *SIMBIOSIS*, vol. 2, no. 2, 2025, doi: 10.30599/simbiosis.v1i2.
- [4] A. C. Amalia, S. Mubarak, and A. Nuraini, "Respons anggrek dendrobium terhadap perbedaan naungan dan aplikasi zat pengatur tumbuh," *Kultivasi*, vol. 21, no. 2, Aug. 2022, doi: 10.24198/kultivasi.v21i2.35029.

- [5] F. Febriansyah and N. Nurmaleni, "Implementasi Algoritma Random Forest Berbasis Machine Learning Untuk Prediksi Klon Kopi Unggul," *Jurnal Komtika (Komputasi dan Informatika)*, vol. 9, no. 2, pp. 230–239, Nov. 2025, doi: 10.31603/komtika.v9i2.15227.
- [6] M. Muin and K. Kusmaladewi, "Efektivitas Peningkatan Pembelajaran Berbasis Kecerdasan Buatan," *Jurnal Studi Guru dan Pembelajaran*, vol. 8, no. 2, pp. 618–631, Jun. 2025, doi: 10.30605/jsdp.8.2.2025.5865.
- [7] A. Nurul Fadillah and A. Selao, "Prediksi Hasil Pertanian Padi Menggunakan Regresi Linear Berbasis Web," 2026, doi: 10.55382/jurnalpustakadata.v6i2.1855.
- [8] F. Purwaningtias, "39 Integrasi Transfer Learning pada Inception v3 untuk Identifikasi Otomatis Jenis Buah Tropis," *JCOSIS (Journal Computer Science and Information Syetem)*, 2025, [Online]. Available: <https://doi.org/10.61567>
- [9] S. Helmiyah and R. Pramestiawan, "Analisis Komparatif Algoritma Machine Learning dengan Metrik Akurasi, Presisi, Recall, dan F1-Score pada Dataset Kacang Kering," *Jurnal Ilmu Komputer dan Teknologi*, vol. 6, no. 3, pp. 152–159, Oct. 2025, doi: 10.35960/ikomti.v6i3.2031.
- [10] O. N. Cahyani and F. Budiman, "Performa Logistic Regression dan Naive Bayes dalam Klasifikasi Berita Hoax di Indonesia," *Edumatic: Jurnal Pendidikan Informatika*, vol. 9, no. 1, pp. 60–68, Apr. 2025, doi: 10.29408/edumatic.v9i1.28987.