

# Analisis Komparatif Algoritma *Machine Learning* dengan Metrik Akurasi, Presisi, *Recall*, dan *F1-Score* pada Dataset Kacang Kering

Siti Helmiyah<sup>1✉</sup>, Rico Pramestiawan<sup>2)</sup>

<sup>1,2)</sup> Program Studi Pendidikan Informatika, Sekolah Tinggi Keguruan dan Ilmu Pendidikan  
Rosalia Lampung, Indonesia

<sup>1)</sup> [shelmiyah.stkiprl@gmail.com](mailto:shelmiyah.stkiprl@gmail.com) ✉

## ABSTRACT

*This study aims to compare the performance of five machine learning algorithms in classifying dry bean varieties as an effort to support quality detection systems for agricultural products. Issues related to authenticity and food safety that frequently occur, such as rice adulteration, highlight the importance of fast and accurate methods for variety identification. The study utilizes the Dry Bean Dataset from the UCI Machine Learning Repository, which consists of 13,611 samples with 16 numerical features and 7 classes of bean varieties. Five algorithms were tested, including K-Nearest Neighbors (KNN), Decision Tree (DT), Support Vector Machine (SVM), Random Forest (RF), and Logistic Regression (LR). The data were divided into 80% for training and 20% for testing, and evaluated using accuracy, precision, recall, and F1-Score metrics. The results show that the SVM algorithm achieved the best performance with an accuracy of 92.43% and an F1-Score of 93.61%, followed by Logistic Regression and Random Forest. The confusion matrix analysis indicates that most varieties were correctly classified, although some misclassifications occurred among classes with similar morphological characteristics such as Dermason, Seker, and Sira. Based on these findings, it can be concluded that selecting the appropriate algorithm is crucial in applying machine learning for agricultural product classification. Evaluation using multiple metrics provides a more comprehensive performance overview compared to relying solely on accuracy. This approach has the potential to support more efficient automation in the identification of agricultural product varieties.*

*Keywords: accuracy, dry bean dataset, f1-score, machine learning algorithms, precision, recall*

## ABSTRAK

Penelitian ini bertujuan membandingkan kinerja dari lima algoritma *machine learning* dalam klasifikasi varietas kacang kering sebagai upaya mendukung sistem deteksi mutu hasil pertanian. Permasalahan keaslian dan keamanan produk pangan yang kerap terjadi, seperti kasus pemalsuan beras, menegaskan pentingnya metode identifikasi varietas yang cepat dan akurat. Penelitian ini menggunakan *Dry Bean Dataset* dari *UCI Machine Learning Repository* yang terdiri atas 13.611 sampel dengan 16 fitur numerik dan 7 kelas varietas kacang. Lima algoritma yang diuji meliputi *K-Nearest Neighbors* (KNN), *Decision Tree* (DT), *Support Vector Machine* (SVM), *Random Forest* (RF), dan *Logistic Regression* (LR). Data dibagi menjadi 80% untuk pelatihan dan 20% untuk pengujian, kemudian dievaluasi menggunakan metrik akurasi, presisi, *recall*, dan *F1-Score*. Hasil penelitian menunjukkan bahwa algoritma SVM menghasilkan performa terbaik dengan akurasi sebesar 92,43% dan *F1-Score* sebesar 93,61%, diikuti oleh *Logistic Regression* dan *Random Forest*. Analisis *confusion matrix* mengindikasikan bahwa sebagian besar varietas dapat diklasifikasikan dengan benar, meskipun masih terdapat kesalahan pada kelas yang memiliki kemiripan morfologi seperti Dermason, Seker, dan Sira. Berdasarkan hasil penelitian ini dapat disimpulkan bahwa pemilihan algoritma yang tepat sangat penting dalam penerapan *machine learning* untuk klasifikasi produk pertanian. Evaluasi dengan berbagai metrik juga memberikan gambaran yang lebih komprehensif dibanding hanya menggunakan akurasi. Pendekatan ini memiliki potensi untuk mendukung otomatisasi identifikasi varietas hasil pertanian secara lebih efisien.

Kata kunci: akurasi, algoritma *machine learning*, *dry bean dataset*, *f1-score*, presisi, *recall*

## I. PENDAHULUAN

Keamanan dan keaslian produk pangan menjadi isu yang semakin krusial dalam sektor pertanian. Pada akhir bulan Juli 2025, publik dikejutkan oleh kasus beras premium yang diduga dioplos dengan beras kualitas rendah. Investigasi yang dilakukan oleh Kementerian Pertanian bersama Bareskrim Polri mengungkap adanya praktik pencampuran beras yang dikemas dengan kemasan beras premium. Sejumlah merek beras populer yang ada dipasaran ikut diperiksa, termasuk produk dari Wilmar Group seperti Sania, Sovia, dan Fortune, sebagaimana diberitakan oleh MetroTV News dan Detik Finance [1], [2]. Laporan Tempo juga menegaskan bahwa beras premium dari beberapa produsen sedang dalam pemeriksaan lebih lanjut [3]. Sementara itu, catatan iNews.id menunjukkan bahwa potensi kerugian dari praktik beras oplosan ini bisa mencapai triliunan rupiah per tahun [4]. Praktik curang semacam ini tidak hanya merugikan konsumen dari sisi mutu dan harga, tetapi juga menurunkan kepercayaan masyarakat terhadap kualitas beras nasional. Fenomena ini menegaskan urgensi keberadaan sistem identifikasi varietas pangan yang cepat, akurat, dan dapat diandalkan untuk mendukung pengawasan mutu.

Upaya mendeteksi kecurangan pangan telah banyak dilakukan di tingkat global. Beberapa studi menunjukkan bahwa kombinasi sensor portabel dengan algoritma *machine learning* dapat mendeteksi pemalsuan beras secara akurat. Misalnya, penggunaan *electronic nose* terbukti efektif dalam membedakan varietas beras [5], sementara penelitian lain memanfaatkan spektroskopi NIR untuk mengidentifikasi pencampuran yang tidak sesuai standar [6]. Lebih lanjut, pendekatan *explainable artificial intelligence* (XAI) mulai diterapkan agar sistem deteksi tidak hanya memberikan hasil prediksi, tetapi juga alasan yang dapat dipahami oleh pengguna [7]. Hal ini menunjukkan bahwa integrasi kecerdasan buatan dengan metode analisis pangan berpotensi besar dalam meminimalisasi praktik kecurangan.

Perkembangan teknologi informasi dan *machine learning* juga telah memberikan dampak signifikan di berbagai bidang, termasuk pertanian. Salah satu tantangan utama dalam pertanian modern adalah klasifikasi varietas tanaman dengan akurasi tinggi. Identifikasi manual membutuhkan keahlian khusus dan sering kali berisiko menghasilkan kesalahan subjektif [8], [9]. Dengan memanfaatkan algoritma *machine learning*, proses klasifikasi dapat dilakukan secara otomatis dengan hasil yang lebih konsisten, cepat, dan efisien, sehingga mendukung peningkatan produktivitas pertanian [10].

Dalam penelitian klasifikasi pertanian, *Dry Bean Dataset* menjadi salah satu dataset *benchmark* yang banyak digunakan. Dataset ini terdiri dari 13.611 sampel dengan 16 fitur numerik yang merepresentasikan tujuh varietas kacang kering [11]. Beberapa penelitian terdahulu telah menggunakan dataset ini dengan metode berbeda, seperti

perbandingan algoritma klasifikasi [12], penerapan metode clustering [13], serta evaluasi kinerja pada data yang tidak seimbang [14].

Sejumlah penelitian lain juga menunjukkan variasi pendekatan terhadap *Dry Bean Dataset*. Koklu dan Ozkan [11] menerapkan algoritma Support Vector Machine dan Random Forest untuk mengklasifikasikan varietas kacang dan memperoleh akurasi di atas 90%. Khan et al. [12] melakukan perbandingan beberapa algoritma klasifikasi multikelas, namun fokus evaluasi masih terbatas pada akurasi. Lee et al [14] mengusulkan penerapan metode *clustering* untuk penanganan data tidak seimbang, sedangkan Chahal dan Kaur [13] mengombinasikan klasifikasi dan *clustering* guna meningkatkan ketepatan prediksi. Kamilaris dan Prenafeta-Boldú [15] menyoroti pemanfaatan *deep learning* pada bidang pertanian, tetapi penelitian tersebut difokuskan pada identifikasi objek seperti daun dan buah, bukan klasifikasi varietas kacang.

Berdasarkan hasil tersebut, dapat diidentifikasi adanya celah penelitian, yaitu belum adanya kajian yang secara komprehensif membandingkan beberapa algoritma *machine learning* menggunakan berbagai metrik evaluasi seperti akurasi, presisi, *recall*, dan *F1-Score*. Penelitian ini berupaya mengisi celah tersebut dengan memberikan analisis komparatif terhadap lima algoritma klasifikasi untuk memperoleh gambaran performa model yang lebih menyeluruh pada kasus klasifikasi varietas kacang kering.

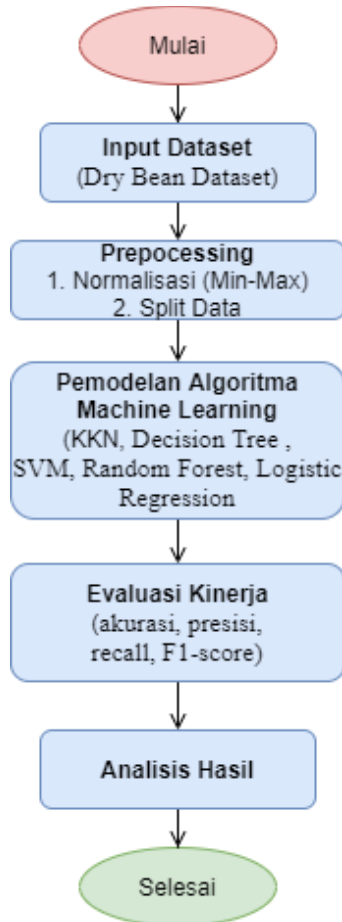
Meskipun hasil tersebut menunjukkan performa yang baik, sebagian besar studi hanya menekankan pada evaluasi berbasis akurasi. Padahal, akurasi saja belum cukup mencerminkan performa model, khususnya pada data multikelas yang kompleks [16]. Oleh karena itu, diperlukan analisis dengan metrik yang lebih lengkap, seperti presisi, *Recall*, dan *F1-Score* [17].

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk membandingkan kinerja lima algoritma *machine learning* yaitu K-Nearest Neighbors (KNN), Decision Tree (DT), Support Vector Machine (SVM), Random Forest (RF), dan Logistic Regression (LR) dalam mengklasifikasikan varietas kacang kering menggunakan *Dry Bean Dataset*. Evaluasi dilakukan dengan pembagian data latih dan data uji sebesar 80:20 menggunakan metrik akurasi, presisi, *Recall*, dan *F1-Score*. Pemilihan kelima algoritma tersebut dilakukan karena masing-masing merepresentasikan pendekatan yang berbeda dalam proses klasifikasi. KNN termasuk metode berbasis instance yang sederhana namun efektif untuk data berukuran kecil hingga menengah. DT merupakan algoritma berbasis pohon keputusan yang mudah diinterpretasikan. RF menggabungkan beberapa pohon keputusan (*ensemble*) sehingga mampu meningkatkan akurasi dan mengurangi risiko *overfitting*. SVM dikenal unggul dalam menangani data berdimensi tinggi dan menghasilkan margin pemisah yang optimal, sedangkan LR berfungsi sebagai model linier yang efisien dan menjadi pembanding dasar dalam klasifikasi biner maupun multikelas. Kombinasi kelima algoritma ini diharapkan dapat memberikan

perbandingan performa yang menyeluruh pada kasus klasifikasi kacang kering. Hasil penelitian diharapkan dapat memberikan rekomendasi algoritma yang tepat untuk klasifikasi produk pertanian sekaligus memperkuat penerapan *machine learning* dalam mendukung pencegahan kecurangan pangan.

## II. METODE PENELITIAN

Bagian ini menjelaskan alur penelitian yang meliputi penggunaan dataset, tahapan *preprocessing*, pemodelan algoritma *machine learning*, evaluasi kinerja dan analisis hasil. Alur penelitian ini dapat dilihat pada Gambar 1.



Gambar 1. Flowchart penelitian klasifikasi biji kacang kering menggunakan algoritma *machine learning*

### A. Dataset

Penelitian ini menggunakan *Dry Bean Dataset* yang tersedia pada *UCI Machine Learning Repository* dan *Kaggle* [11]. Dataset ini terdiri atas 13.611 sampel biji kacang kering yang mencakup tujuh kelas varietas kacang yaitu Barbunya, Bombay, Cali, Dermason, Horoz, Seker, dan Sira. Masing-masing sampel direpresentasikan oleh 16 fitur numerik yang menggambarkan karakteristik morfologi biji kacang hasil ekstraksi citra, seperti panjang, lebar, area, dan keliling. Tabel 1 berikut menyajikan daftar fitur pada *Dry Bean Dataset* beserta tipe data dan kategorinya dan Tabel 2 menyajikan cuplikan *Dry Bean*.

Tabel 1. Daftar Fitur pada *Dry Bean Dataset*

No	Nama Fitur	Tipe Data	Kategori
1	Area	Numerik (float)	Non-Class
2	Perimeter	Numerik (float)	Non-Class
3	MajorAxisLength	Numerik (float)	Non-Class
4	MinorAxisLength	Numerik (float)	Non-Class
5	AspectRatio	Numerik (float)	Non-Class
6	Eccentricity	Numerik (float)	Non-Class
7	ConvexArea	Numerik (float)	Non-Class
8	EquivalentDiameter	Numerik (float)	Non-Class
9	Extent	Numerik (float)	Non-Class
10	Solidity	Numerik (float)	Non-Class
11	Roundness	Numerik (float)	Non-Class
12	Compactness	Numerik (float)	Non-Class
13	ShapeFactor1	Numerik (float)	Non-Class
14	ShapeFactor2	Numerik (float)	Non-Class
15	ShapeFactor3	Numerik (float)	Non-Class
16	ShapeFactor4	Numerik (float)	Non-Class
17	Class	Kategorikal (string)	Class

Tabel 2. Cuplikan Dataset *Dry Bean*

Area	Perimeter	Major Axis Length	....	Class
28395	610,291	208,1781167		Seker
28734	638,018	200,5247957		Seker
29380	624,11	212,8261299		Barbunya
42012	809,77	266,7003964		Barbunya
43122	838,186	296,4045889		Bombay
43145	799,426	273,4029446		Bombay
114004	1279,356	451,3612558		Cali
117034	1265,926	425,9237875		Cali
49245	822,642	301,392791		Horoz
49882	891,505	357,1890036		Horoz
33006	710,496	283,0203847		Sira
33263	719,325	271,3390351		Sira
20420	524,932	183,601165		Dermason
20464	528,408	191,249312		Dermason

*Dry Bean Dataset* telah banyak dimanfaatkan dalam penelitian klasifikasi pertanian. Beberapa studi menunjukkan perbandingan algoritma klasifikasi [12], penerapan metode *clustering* [13], serta evaluasi kinerja pada data yang tidak seimbang [14]. Temuan-temuan ini menegaskan bahwa peran *Dry Bean Dataset* sebagai *benchmark* yang relevan untuk mengevaluasi performa algoritma *machine learning* pada data multikelas.

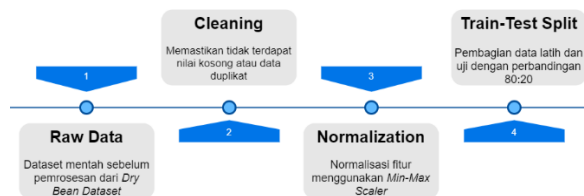
### B. Preprocessing Data

Sebelum dilakukan pemodelan, dataset melewati tahap pra-pemrosesan. Tahapan awal meliputi *data cleaning* untuk memastikan tidak terdapat nilai kosong (*missing values*) atau data duplikat yang dapat memengaruhi hasil pelatihan model. Seluruh fitur numerik dinormalisasi menggunakan *Min-Max Scaler* ke dalam rentang [0–1] untuk mengurangi bias akibat perbedaan skala antar fitur. Normalisasi penting dilakukan karena algoritma seperti KNN dan SVM sensitif terhadap skala fitur [18]. Selanjutnya, dataset dibagi menjadi 80% data latih (10.889 sampel) dan 20% data uji (2.722 sampel). Proses pembagian dilakukan secara acak dengan penetapan *random\_state*

untuk menjaga konsistensi hasil, serta menggunakan teknik stratifikasi agar distribusi kelas tetap seimbang.

Rasio 80:20 dipilih karena secara umum memberikan keseimbangan yang baik antara jumlah data untuk pelatihan dan proporsi data untuk pengujian. Rasio ini banyak digunakan dalam penelitian klasifikasi karena memastikan model memperoleh cukup data untuk belajar tanpa mengurangi kualitas pengujian [19], [20]. Percobaan awal dengan rasio 70:30 menunjukkan bahwa peningkatan proporsi data uji tidak memberikan perbedaan signifikan terhadap akurasi model, sementara penggunaan data latih yang lebih sedikit cenderung menurunkan stabilitas hasil. Oleh karena itu, rasio 80:20 dipertahankan sebagai konfigurasi terbaik dalam penelitian ini.

Alur proses *preprocessing* data ditunjukkan pada Gambar 2, yang menggambarkan tahapan mulai dari data mentah hingga terbentuknya data latih dan data uji yang siap digunakan dalam pemodelan. Tahapan ini sejalan dengan metodologi umum dalam penelitian klasifikasi multikelas berbasis *machine learning* [21].



Gambar 2. Alur *Preprocessing* Data

### C. Algoritma *Machine learning*

Penelitian ini membandingkan kinerja lima algoritma klasifikasi yang telah banyak digunakan dalam literatur, yaitu:

#### 1) K-Nearest Neighbors (KNN)

KNN bekerja berdasarkan prinsip kedekatan antar data [22], dengan menghitung jarak Euclidean antara sampel uji dan data latih. Pada penelitian ini, nilai  $k$  ditentukan melalui proses uji coba (*experimentation*), dan klasifikasi dilakukan berdasarkan kelas mayoritas dari  $k$  tetangga terdekat. Algoritma ini sederhana dan efektif untuk data berdimensi sedang, namun sensitif terhadap skala fitur sehingga tahap normalisasi menjadi penting untuk menjaga akurasi.

#### 2) Decision Tree (DT)

DT membangun struktur pohon keputusan yang memisahkan data berdasarkan fitur dengan *information gain* tertinggi [23]. Dalam konteks klasifikasi kacang kering, setiap simpul pohon memisahkan data berdasarkan atribut morfologi seperti panjang, lebar, atau area biji. Model ini mudah diinterpretasikan, namun dapat mengalami *overfitting* apabila kedalaman pohon tidak diatur dengan tepat.

#### 3) Support Vector Machine (SVM)

SVM mencari *hyperplane* optimal yang memisahkan kelas pada ruang berdimensi tinggi

[24]. Pada penelitian ini, digunakan fungsi kernel linier karena hubungan antar fitur pada *Dry Bean Dataset* cenderung dapat dipisahkan secara linier. SVM mampu menghasilkan margin pemisah yang maksimal antar kelas, sehingga memberikan akurasi tinggi terutama pada data multikelas yang memiliki fitur kompleks.

#### 4) Random Forest (RF)

RF merupakan metode *ensemble* yang menggabungkan beberapa pohon keputusan untuk meningkatkan akurasi dan kestabilan model [25]. Setiap pohon dilatih menggunakan subset data yang dipilih secara acak, kemudian hasil prediksi digabungkan melalui voting mayoritas. Pendekatan ini mampu mengurangi risiko *overfitting* yang sering terjadi pada satu pohon keputusan dan memberikan performa yang lebih stabil.

#### 5) Logistic Regression (LR)

LR digunakan sebagai model linier yang memprediksi probabilitas kelas menggunakan fungsi sigmoid [26]. Dalam penelitian ini, LR berfungsi sebagai pembanding dasar untuk menilai performa model linier terhadap algoritma non-linier lainnya. Meskipun sederhana, LR tetap relevan karena efisien dan mampu memberikan interpretasi yang jelas terhadap hubungan antar fitur.

Pemilihan kelima algoritma tersebut merepresentasikan variasi pendekatan dalam proses klasifikasi, mulai dari metode berbasis instance (KNN), pohon keputusan (DT), *ensemble* (RF), margin-based (SVM), hingga model linier (LR). Keragaman ini memberikan dasar yang kuat untuk analisis komparatif dalam penelitian ini. Selain itu, algoritma-algoritma tersebut telah terbukti efektif dalam berbagai studi klasifikasi pertanian [15], sehingga relevan diterapkan untuk kasus klasifikasi varietas kacang kering. Seluruh implementasi dilakukan menggunakan pustaka *Scikit-learn* di Python, guna memastikan konsistensi hasil dan memudahkan replikasi penelitian oleh peneliti lain.

### D. Evaluasi Kinerja

Evaluasi dilakukan menggunakan empat metrik utama: akurasi, presisi, *Recall*, dan *F1-Score*.

- 1) Akurasi digunakan untuk mengukur proporsi prediksi yang benar terhadap keseluruhan data uji.
- 2) Presisi digunakan untuk menilai ketepatan prediksi positif dibandingkan semua prediksi positif.
- 3) *Recall* digunakan untuk mengukur sensitivitas model dalam mengenali data positif pada tiap kelas.
- 4) *F1-Score* digunakan untuk harmonisasi antara presisi dan *Recall*, yang berguna pada data multikelas tidak seimbang [27].

Penggunaan metrik selain akurasi dianggap penting karena akurasi saja seringkali tidak cukup untuk menggambarkan performa model secara menyeluruh, terutama dalam kasus klasifikasi dengan distribusi kelas yang tidak seimbang [28]. Untuk memperdalam analisis, ditampilkan *confusion matrix* pada algoritma dengan performa terbaik, sehingga dapat diketahui distribusi prediksi benar dan salah antar kelas serta pola kesalahan klasifikasi.

### III. HASIL DAN PEMBAHASAN

Bagian ini memaparkan hasil eksperimen dan analisis komparatif terhadap lima algoritma *machine learning* yang digunakan dalam penelitian. Uji coba dilakukan pada data uji dengan porsi 20% dari total dataset, menggunakan metrik akurasi, presisi, *recall*, dan *F1-score*. Hasil yang diperoleh kemudian dianalisis untuk mengetahui kelebihan serta keterbatasan masing-masing algoritma, sekaligus membandingkannya dengan temuan penelitian terdahulu.

#### A. Hasil Penelitian

Bagian ini memaparkan hasil eksperimen dan analisis komparatif terhadap lima algoritma *machine learning* yang digunakan dalam penelitian. Uji coba dilakukan pada data uji dengan porsi 20% dari total dataset, menggunakan metrik akurasi, presisi, *recall*, dan *F1-score*. Hasil yang diperoleh kemudian dianalisis untuk mengetahui kelebihan serta keterbatasan masing-masing algoritma, sekaligus membandingkannya dengan temuan penelitian terdahulu.

Hasil pengujian yang dilakukan pada 20% data uji yaitu sebanyak 2.722 sampel untuk mengevaluasi performa lima algoritma *machine learning* menggunakan metrik akurasi, presisi, *recall*, dan *F1-Score* dapat dilihat pada Tabel 3.

Pemilihan rasio pembagian data latih dan data uji sebesar 80:20 didasarkan pada pertimbangan metodologis untuk menjaga keseimbangan antara jumlah data pelatihan yang memadai dan data pengujian yang representatif. Rasio ini banyak digunakan dalam penelitian klasifikasi karena dinilai mampu menghasilkan model yang stabil dan akurat [19], [20]. Alternatif seperti rasio 70:30 memang umum digunakan, namun pada dataset besar seperti *Dry Bean Dataset*, proporsi data latih yang lebih tinggi dianggap lebih optimal karena memungkinkan model mempelajari pola dengan lebih baik tanpa mengurangi kualitas evaluasi. Oleh karena itu, rasio 80:20 dipertahankan sebagai konfigurasi terbaik dalam penelitian ini.

Tabel 3. Hasil evaluasi model pada *Dry Bean Dataset*

Algoritma	Akurasi	Presisi (macro)	Recall (macro)	F1-Score (macro)
SVM	92,43 %	93,72 %	93,50 %	93,61 %
Logistic Regression	92,07 %	93,49 %	93,14 %	93,29 %

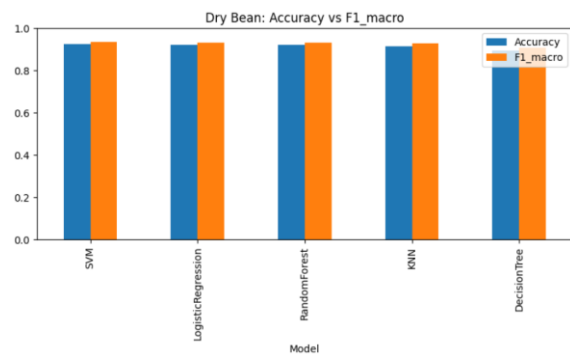
Random Forest	91,99 %	93,39 %	93,06 %	93,21 %
KNN	91,37 %	92,90 %	92,45 %	92,65 %
Decision Tree	89,20 %	90,75 %	90,90 %	90,81 %

Evaluasi performa model dilakukan menggunakan pustaka *Scikit-learn* pada bahasa pemrograman Python. Proses pelatihan dan pengujian dilakukan secara terpisah, di mana model dilatih menggunakan 80% data latih dan dievaluasi pada 20% data uji. Setelah proses pelatihan, dilakukan prediksi terhadap data uji untuk memperoleh nilai akurasi, presisi, *recall*, dan *F1-Score* dengan memanfaatkan fungsi *classification\_report()* dari modul *sklearn.metrics*.

Selanjutnya, hasil prediksi divisualisasikan dalam bentuk *confusion matrix* menggunakan fungsi *confusion\_matrix()* dan *ConfusionMatrixDisplay()*. Visualisasi dibuat dengan pustaka *Matplotlib* untuk menampilkan pola kesalahan klasifikasi tiap kelas secara visual. Pendekatan ini memberikan gambaran yang lebih komprehensif terhadap performa model karena memungkinkan analisis terhadap distribusi prediksi benar (diagonal) dan salah (non-diagonal).

Berdasarkan Tabel 3, algoritma Support Vector Machine (SVM) menunjukkan performa terbaik dengan akurasi 92,43% dan *F1-Score* 93,61%. Algoritma Logistic Regression dan Random Forest berada di peringkat berikutnya dengan akurasi 92,07% dan 91,99%. Sementara KNN memperoleh akurasi 91,36%, dan Decision Tree memiliki performa terendah dengan akurasi 89,20%.

Hasil evaluasi dari pengujian data uji pada *Dry Bean Dataset* grafik perbandingan akurasi dan *F1-Score* dapat dilihat pada Gambar 3.



Gambar 3. Grafik perbandingan akurasi dan *F1-Score* lima algoritma *machine learning*

Gambar 3 menunjukkan bahwa algoritma SVM memiliki keunggulan paling tinggi, selanjutnya diikuti algoritma Logistic Regression dan algoritma Random Forest. Sementara algoritma KNN dan algoritma Decision Tree menunjukkan performa yang lebih rendah secara konsisten pada kedua metrik.

Secara teoretis, keunggulan algoritma SVM disebabkan oleh kemampuannya membentuk *hyperplane* optimal yang memisahkan setiap kelas dengan margin maksimum, sehingga menghasilkan batas keputusan yang lebih stabil dan generalisasi yang



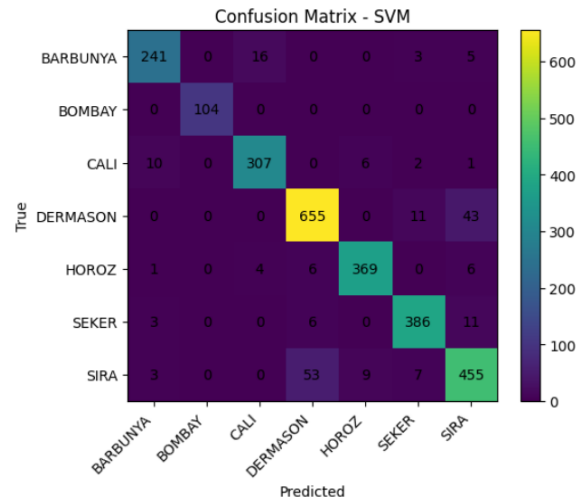
lebih baik pada data baru [24]. Pendekatan ini sesuai dengan prinsip *structural risk minimization*, di mana SVM tidak hanya meminimalkan kesalahan pada data latih, tetapi juga mengontrol kompleksitas model agar tidak terjadi *overfitting*. Kemampuan ini membuat SVM efektif untuk data berdimensi tinggi seperti *Dry Bean Dataset* yang memiliki 16 fitur numerik. Sebaliknya, algoritma Decision Tree (DT) menunjukkan performa yang lebih rendah karena mekanisme pembelajarannya bergantung pada pemisahan berbasis ambang nilai fitur (*threshold split*), yang rentan terhadap variasi kecil pada data [23]. DT juga cenderung membentuk model yang terlalu spesifik terhadap data latih, terutama pada dataset besar dengan banyak fitur yang saling berkorelasi, sehingga menyebabkan *overfitting* [25]. Hal ini menjelaskan mengapa SVM mampu memberikan hasil yang paling konsisten, sedangkan DT menghasilkan akurasi dan *F1-Score* terendah pada pengujian.

Berdasarkan hasil pada Tabel 3 dan Gambar 3, algoritma Support Vector Machine (SVM) menghasilkan performa terbaik dengan akurasi 92,43% dan *F1-Score* 93,61%. Keunggulan ini disebabkan oleh kemampuan SVM dalam membentuk *hyperplane* optimal yang memisahkan setiap kelas dengan margin maksimum. Dengan 16 fitur numerik pada *Dry Bean Dataset* yang menggambarkan karakteristik morfologi biji kacang (panjang, lebar, area, perimeter, dan sebagainya), SVM dapat bekerja efektif pada data berdimensi tinggi dan tetap stabil meskipun terdapat variasi antar kelas. Penggunaan kernel nonlinier membantu model ini menemukan batas keputusan yang lebih akurat di antara varietas yang memiliki kemiripan bentuk.

Sebaliknya, algoritma Decision Tree (DT) menunjukkan performa terendah dengan akurasi 89,20% dan *F1-Score* 90,81%. Kinerja rendah ini disebabkan oleh kecenderungan DT untuk *overfitting* terhadap data latih, terutama pada dataset dengan banyak fitur numerik yang saling berkorelasi. Selain itu, metode pemisahan berbasis nilai ambang sederhana (*threshold*) membuat DT kurang mampu menangkap hubungan nonlinier antar fitur, sehingga model sulit membedakan varietas kacang yang memiliki bentuk serupa. Kondisi ini menjelaskan mengapa DT menghasilkan akurasi dan *F1-Score* yang lebih rendah dibandingkan algoritma lainnya.

Dengan demikian, dapat disimpulkan bahwa SVM unggul karena kemampuannya memaksimalkan margin dan menangani data multidimensi secara efektif, sedangkan DT kurang optimal karena sensitivitasnya terhadap variasi fitur dan kecenderungannya menghasilkan model yang tidak stabil.

Untuk memperkuat temuan tersebut, analisis selanjutnya disajikan melalui visualisasi confusion matrix untuk algoritma dengan performa terbaik, yaitu SVM. *Confusion matrix* algoritma SVM ini dapat dilihat pada Gambar 4.



Gambar 4. *Confusion matrix* untuk SVM sebagai model terbaik

Gambar 4 menunjukkan bahwa sebagian besar kelas kacang dapat diprediksi dengan benar, meskipun masih terdapat kesalahan klasifikasi pada varietas yang memiliki kemiripan morfologi, seperti Dermason, Seker, dan Sira.

*Confusion matrix* pada algoritma SVM sebagai model terbaik ditunjukkan pada Gambar 4. Visualisasi ini memperlihatkan bahwa sebagian besar kelas berhasil diklasifikasikan dengan benar. Namun, masih terdapat kesalahan klasifikasi pada varietas dengan morfologi mirip seperti Dermason, Seker, dan Sira.

## B. Pembahasan

Hasil penelitian ini menegaskan bahwa SVM unggul dalam klasifikasi *Dry Bean Dataset*, terutama berkat kemampuannya mengelola data berdimensi tinggi dan multikelas [22]. Logistic Regression, meskipun merupakan model linier sederhana, terbukti kompetitif dengan akurasi di atas 92%. Random Forest juga memberikan hasil stabil dengan akurasi tinggi, meskipun membutuhkan waktu latih lebih lama akibat proses *ensemble*.

Kinerja KNN relatif baik, tetapi sensitif terhadap distribusi data dan pemilihan parameter jumlah tetangga. Decision Tree mencatat akurasi terendah, kemungkinan karena kecenderungan *overfitting* pada data latih sehingga tidak mampu menangkap kompleksitas data dengan baik.

Penggunaan metrik multi-evaluasi (akurasi, presisi, *Recall*, dan *F1-Score*) terbukti lebih informatif daripada hanya mengandalkan akurasi [16], [17]. Sebagai contoh, selisih akurasi antar-algoritma relatif kecil, namun nilai *F1-Score* memberikan gambaran keseimbangan performa antar kelas yang lebih jelas.

Temuan ini konsisten dengan penelitian terdahulu yang menekankan efektivitas SVM pada dataset *Dry Bean* [12]. Penelitian lain juga melaporkan performa kompetitif algoritma serupa pada konteks klasifikasi multikelas [13], [14].

Lebih jauh, hasil ini menunjukkan bahwa pemilihan algoritma yang tepat berpotensi mendukung pengembangan sistem otomatisasi identifikasi varietas pangan. Studi terkait kasus beras oplosan menyoroiti urgensi deteksi mutu pangan yang lebih akurat [1], [2], sementara penelitian lain menegaskan bahwa praktik pengoplosan beras premium berdampak signifikan terhadap kepercayaan konsumen [3], [4].

Meskipun hasil penelitian menunjukkan bahwa SVM memberikan performa terbaik pada *Dry Bean Dataset*, penerapannya di lapangan tetap memiliki keterbatasan. Dataset yang digunakan relatif bersih dan seimbang, sedangkan data nyata sering kali menghadapi masalah noise, ketidakseimbangan kelas, atau variasi kualitas citra. Kondisi ini berpotensi menurunkan performa model jika langsung diterapkan. Oleh karena itu, penelitian selanjutnya dapat mempertimbangkan penggunaan teknik data augmentation, metode penanganan *imbalanced* data, atau pendekatan *deep learning* untuk meningkatkan generalisasi model.

Dari sisi praktis, hasil ini menegaskan bahwa algoritma *machine learning* yang tepat dapat mendukung sistem otomatisasi identifikasi varietas pangan secara lebih cepat, akurat, dan konsisten. Implementasi sistem semacam ini dapat membantu pengawasan mutu pertanian sekaligus mencegah praktik kecurangan pangan di Indonesia.

#### IV. KESIMPULAN

Penelitian ini membandingkan kinerja lima algoritma *machine learning* K-Nearest Neighbors (KNN), Decision Tree (DT), Support Vector Machine (SVM), Random Forest (RF), dan Logistic Regression (LR) dalam klasifikasi varietas kacang kering menggunakan *Dry Bean Dataset*. Hasil eksperimen menunjukkan bahwa SVM memberikan performa terbaik dengan akurasi 92,43% dan *F1-Score* 93,61%, diikuti Logistic Regression dan Random Forest dengan selisih tipis. KNN memberikan hasil yang cukup baik, sementara Decision Tree menunjukkan performa terendah akibat kecenderungan *overfitting*.

Penggunaan metrik evaluasi yang lebih komprehensif, yaitu akurasi, presisi, *Recall*, dan *F1-Score*, terbukti memberikan gambaran yang lebih lengkap dibandingkan hanya mengandalkan akurasi. Hal ini penting terutama pada data multikelas, di mana distribusi performa antar kelas perlu dianalisis secara lebih mendalam.

Secara praktis, hasil penelitian ini menegaskan bahwa pemilihan algoritma yang tepat sangat berpengaruh dalam pengembangan sistem identifikasi varietas pangan. Implementasi model klasifikasi berbasis *machine learning* dapat mendukung upaya peningkatan mutu pertanian sekaligus membantu pencegahan praktik kecurangan pangan.

Untuk penelitian selanjutnya, disarankan melakukan pengujian pada dataset dengan kondisi yang lebih kompleks, seperti data tidak seimbang atau mengandung *noise*, serta mengeksplorasi metode

lanjutan seperti *deep learning* atau *ensemble learning* tingkat lanjut guna meningkatkan generalisasi model.

#### REFERENSI

- [1] S. Y. Hukmana, "Total 22 Saksi Diperiksa Kasus Beras Premium Oplosan, Termasuk 6 Produsen." Accessed: Sep. 28, 2025. [Online]. Available: <https://www.metrotvnews.com/read/bw6CgBxG-total-22-saksi-diperiksa-kasus-beras-premium-oplosan-termasuk-6-produsen>
- [2] A. Damayanti, "Daftar Merek Beras yang Diduga Dioplos, Ada Sania hingga Ayana Baca artikel detikfinance, "Daftar Merek Beras yang Diduga Dioplos, Ada Sania hingga Ayana." Accessed: Sep. 28, 2025. [Online]. Available: <https://finance.detik.com/berita-ekonomi-bisnis/d-8010316/daftar-merek-beras-yang-diduga-dioplos-ada-sania-hingga-ayana>
- [3] Tempo, "Beras Biasa Dijual Premium, Bareskrim Periksa 4 Produsen Beras." Accessed: Sep. 28, 2025. [Online]. Available: <https://www.tempo.co/hukum/beras-biasa-dijual-premium-bareskrim-periksa-4-produsen-beras--1995462>
- [4] A. Chan, "Bareskrim Polri Bongkar Praktik Curang: 201 Ton Beras Oplosan Disita, Kerugian Capai Rp99 Triliun." Accessed: Sep. 28, 2025. [Online]. Available: <https://sorongraya.inews.id/read/621994/bareskrim-polri-bongkar-praktik-curang-201-ton-beras-oplosan-disita-kerugian-capai-rp99-triliun>
- [5] F. Falcone, A. Aznan, C. Gonzalez Viejo, A. Pang, and S. Fuentes, "Rapid Detection of Rice Adulteration Using a Low-Cost Electronic Nose and Machine Learning Modelling," *Engineering Proceedings 2022*, Vol. 27, Page 1, vol. 27, no. 1, p. 1, Nov. 2022, doi: 10.3390/ECSA-9-13291.
- [6] N. Fazeli Burestan, A. H. Afkari Sayyah, and E. Taghinezhad, "Prediction of some quality properties of rice and its flour by near-infrared spectroscopy (NIRS) analysis," *Food Sci Nutr*, vol. 9, no. 2, pp. 1099–1105, Feb. 2021, doi: 10.1002/FSN3.2086.
- [7] M. T. Ribeiro, S. Singh, and C. Guestrin, "Model-Agnostic Interpretability of Machine Learning," *arXiv.org*, 2016.
- [8] A. Sharma, A. Jain, P. Gupta, and V. Chowdary, "Machine Learning Applications for Precision Agriculture: A Comprehensive Review," *IEEE Access*, vol. 9, pp. 4843–4873, 2021, doi: 10.1109/ACCESS.2020.3048415.
- [9] J. A. Wani, S. Sharma, M. Muzamil, S. Ahmed, S. Sharma, and S. Singh, "Machine Learning and Deep Learning Based Computational Techniques in Automatic Agricultural Diseases Detection: Methodologies, Applications, and Challenges," *Archives of Computational Methods in Engineering*, vol. 29, no. 1, pp. 641–677, Jan. 2022, doi: 10.1007/S11831-021-09588-5/METRICS.
- [10] A. Koirala, K. B. Walsh, Z. Wang, and C. McCarthy, "Deep learning – Method overview and review of use for fruit detection and yield estimation," *Comput Electron Agric*, vol. 162, pp. 219–234, Jul. 2019, doi: 10.1016/J.COMPAG.2019.04.017.
- [11] M. Koklu and I. A. Ozkan, "Multiclass classification of dry beans using computer vision and machine learning techniques," *Comput Electron Agric*, vol. 174, p. 105507, Jul. 2020, doi: 10.1016/J.COMPAG.2020.105507.
- [12] M. Salaudin Khan *et al.*, "Comparison of multiclass classification techniques using dry bean dataset," *International Journal of Cognitive Computing in Engineering*, vol. 4, pp. 6–20, Jun. 2023, doi: 10.1016/J.IJCC.2023.01.002.
- [13] J. K. Chahal and A. Kaur, "A Hybrid Approach based on Classification and Clustering for Intrusion Detection System," *International Journal of Mathematical Sciences and Computing*, vol. 2, no. 4, pp. 34–40, Nov. 2016, doi: 10.5815/IJMSC.2016.04.04.
- [14] C. Y. Lee, W. Wang, and J. Q. Huang, "Clustering and classification for dry bean feature imbalanced data," *Sci Rep*, vol. 14, no. 1, pp. 1–19, Dec. 2024, doi: 10.1038/S41598-024-82253-6;SUBJMETA.

- [15] T. Saito and M. Rehmsmeier, "The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets," *PLoS One*, vol. 10, no. 3, p. e0118432, Mar. 2015, doi: 10.1371/JOURNAL.PONE.0118432.
- [16] D. Chicco and G. Jurman, "The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation," *BMC Genomics*, vol. 21, no. 1, pp. 1–13, Jan. 2020, doi: 10.1186/S12864-019-6413-7/TABLES/5.
- [17] S. Aksoy and R. M. Haralick, "Feature normalization and likelihood-based similarity measures for image retrieval," *Pattern Recognit Lett*, vol. 22, no. 5, pp. 563–582, Apr. 2001, doi: 10.1016/S0167-8655(00)00112-4.
- [18] F. Pedregosa *et al.*, "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, Jan. 2012, Accessed: Oct. 28, 2025. [Online]. Available: <https://arxiv.org/pdf/1201.0490>
- [19] J. Han, M. Kamber, and J. Pei, "Data Mining. Concepts and Techniques, 3rd Edition (The Morgan Kaufmann Series in Data Management Systems)," 2011.
- [20] S. Raschka and V. Mirjalili, "Python Machine Learning Third Edition Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow 2," 2019, Accessed: Oct. 28, 2025. [Online]. Available: [www.packt.com](http://www.packt.com)
- [21] T. M. Cover and P. E. Hart, "Nearest Neighbor Pattern Classification," *IEEE Trans InfTheory*, vol. 13, no. 1, pp. 21–27, 1967, doi: 10.1109/TIT.1967.1053964.
- [22] J. R. Quinlan, "Induction of decision trees," *Machine Learning 1986 1:1*, vol. 1, no. 1, pp. 81–106, Mar. 1986, doi: 10.1007/BF00116251.
- [23] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning 1995 20:3*, vol. 20, no. 3, pp. 273–297, Sep. 1995, doi: 10.1007/BF00994018.
- [24] L. Breiman, "Random forests," *Mach Learn*, vol. 45, no. 1, pp. 5–32, Oct. 2001, doi: 10.1023/A:1010933404324/METRICS.
- [25] D. W. Hosmer and S. Lemeshow, "Applied Logistic Regression," *Applied Logistic Regression*, Sep. 2000, doi: 10.1002/0471722146.
- [26] A. Kamilaris and F. X. Prenafeta-Boldú, "Deep learning in agriculture: A survey," *Comput Electron Agric*, vol. 147, pp. 70–90, Apr. 2018, doi: 10.1016/J.COMPAG.2018.02.016.
- [27] "Encyclopedia Of Machine Learning And Data Mining." Accessed: Sep. 30, 2025. [Online]. Available: <https://pdf-up.com/download/encyclopedia-of-machine-learning-and-data-mining-4975350>
- [28] P. Branco, L. Torgo, and R. P. Ribeiro, "A Survey of Predictive Modeling on Imbalanced Domains," *ACM Computing Surveys (CSUR)*, vol. 49, no. 2, Aug. 2016, doi: 10.1145/2907070.