

Analisis Prediktif Tingkat Kematangan Alpukat Menggunakan Algoritma Logistic Regression

Natasya Aulia¹, Iqbal², Hannan Asrawi³

^{1,2,3}Universitas Almuslim

ARTICLE INFO

Article history:

Received : 26 Februari 2026

Revised : 03 Maret 2026

Accepted : 03 Maret 2026

Keywords:

Avocado Ripeness, Machine Learning, Classification, Logistic Regression



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

Kata kunci:

Kematangan Alpukat, Machine Learning, Klasifikasi, Logistic Regression

Corresponding Author:

Natasya Aulia

Universitas Almuslim

Email:

tasyaalya0806@gmail.com

ABSTRACT

[Predictive Analysis of Avocado Ripeness Level Using Logistic Regression Algorithm] Determining the ripeness level of avocado fruit is an important factor in distribution, marketing, and consumption processes. Conventional ripeness assessment is often subjective and dependent on human experience, which can lead to inconsistent results. This study aims to develop an avocado ripeness prediction system using the Logistic Regression algorithm based on physical and visual fruit characteristics. The dataset consists of 1,250 avocado samples with features including firmness, color attributes, tapping sound, weight, and fruit size. Data preprocessing involved cleaning, normalization of numerical features using StandardScaler, and categorical feature transformation using one-hot encoding. The experimental results show that the proposed model achieved an accuracy of approximately 77% in classifying avocado ripeness into ripe and unripe categories, indicating that Logistic Regression is a lightweight and efficient approach for numerical-based ripeness prediction systems.

ABSTRAK

Penentuan tingkat kematangan buah alpukat merupakan faktor penting dalam proses distribusi, pemasaran, dan konsumsi. Penilaian kematangan secara konvensional masih bersifat subjektif dan bergantung pada pengalaman manusia sehingga dapat menghasilkan penilaian yang tidak konsisten. Penelitian ini bertujuan untuk mengembangkan sistem prediksi tingkat kematangan buah alpukat menggunakan algoritma Logistic Regression berdasarkan karakteristik fisik dan visual buah. Dataset yang digunakan terdiri dari 1.250 data alpukat dengan fitur firmness, warna, suara ketukan, berat, dan ukuran buah. Tahapan preprocessing meliputi pembersihan data, normalisasi fitur numerik menggunakan StandardScaler, serta transformasi fitur kategorikal dengan one-hot encoding. Hasil pengujian menunjukkan bahwa model mencapai akurasi sekitar 77% dalam mengklasifikasikan tingkat kematangan alpukat, sehingga algoritma Logistic Regression dapat digunakan sebagai metode yang efisien dan mudah diimplementasikan.

PENDAHULUAN

Perkembangan teknologi pengolahan data dan *machine learning* telah banyak dimanfaatkan dalam pengembangan sistem klasifikasi dan prediksi secara otomatis pada berbagai bidang, termasuk pertanian dan hortikultura [1],[2],[3]. Pendekatan berbasis data

numerik dan visual digunakan untuk mengidentifikasi kondisi objek seperti buah, sayuran, dan produk pertanian lainnya dengan tujuan meningkatkan akurasi penilaian serta mengurangi subjektivitas pengamatan manual [4][5],[6]. Beberapa penelitian menunjukkan bahwa metode klasifikasi berbasis *machine learning* mampu memberikan hasil yang lebih konsisten dibandingkan metode konvensional yang bergantung pada pengalaman manusia [7],[8], [9].

Dalam bidang pertanian, khususnya pada komoditas buah-buahan, penentuan tingkat kematangan merupakan aspek penting yang memengaruhi kualitas, daya simpan, serta nilai jual produk [10], [11]. Buah alpukat merupakan salah satu komoditas hortikultura yang banyak dikonsumsi dan memiliki tingkat kematangan yang sulit ditentukan secara visual [12], [13]. Perbedaan tingkat kematangan alpukat sering kali tidak terlihat jelas dari warna kulit saja, sehingga metode manual seperti perabaan atau penilaian subjektif sering menghasilkan kesalahan klasifikasi [14],[15]. Kondisi ini mendorong perlunya sistem prediksi kematangan buah alpukat yang lebih objektif dan berbasis data [16],[17].

Berbagai penelitian sebelumnya telah menerapkan metode *machine learning* untuk memprediksi tingkat kematangan buah menggunakan fitur fisik dan visual, seperti warna, tekstur, kekerasan, dan berat buah [18],[19]. Algoritma yang umum digunakan antara lain K-Nearest Neighbor (KNN), Support Vector Machine (SVM), Artificial Neural Network (ANN), dan Convolutional Neural Network (CNN) [20],[21], [22],[23]. Metode berbasis *deep learning*, khususnya CNN, terbukti mampu menghasilkan akurasi yang tinggi, namun umumnya memerlukan dataset berukuran besar serta sumber daya komputasi yang tinggi [24],[25]. Kondisi tersebut menjadikan pendekatan ini kurang efisien untuk sistem sederhana atau aplikasi dengan keterbatasan perangkat keras [26].

Sebagai alternatif, algoritma klasifikasi yang lebih ringan seperti Logistic Regression banyak digunakan pada berbagai permasalahan klasifikasi data karena memiliki kompleksitas komputasi yang rendah dan mudah diimplementasikan [27]. Algoritma ini telah diterapkan pada berbagai studi, termasuk prediksi risiko penyakit, klasifikasi data pertanian, serta sistem pendukung keputusan berbasis data numerik dengan hasil yang cukup baik [28]. Dengan proses pelatihan yang relatif cepat dan interpretasi model yang sederhana, Logistic Regression menjadi pilihan yang sesuai untuk pengembangan sistem prediksi berbasis data numerik [29].

Berdasarkan uraian tersebut, penelitian ini mengembangkan sistem prediksi tingkat kematangan buah alpukat menggunakan algoritma Logistic Regression berdasarkan karakteristik fisik dan visual buah, meliputi kekerasan, warna, suara ketukan, berat, dan ukuran buah. Penelitian ini bertujuan menghasilkan sistem prediksi yang ringan, efisien, dan mudah diimplementasikan, namun tetap memiliki tingkat akurasi yang baik. Diharapkan hasil penelitian ini dapat menjadi alternatif solusi dalam penentuan kematangan buah alpukat secara objektif dan konsisten serta dapat dikembangkan lebih lanjut pada sistem pendukung keputusan di bidang pertanian.

METODE

Penelitian ini disusun melalui beberapa tahapan terstruktur, dimulai dari pengumpulan data alpukat, preprocessing data, pelatihan model, hingga evaluasi kinerja sistem. Dataset yang digunakan terdiri dari 1.250 data alpukat dengan label tingkat kematangan.

Tahap preprocessing meliputi pembersihan data, normalisasi fitur numerik menggunakan StandardScaler, serta transformasi fitur kategorikal warna kulit alpukat menggunakan teknik one-hot encoding. Proses ini bertujuan menyamakan skala fitur agar tidak memengaruhi pembentukan model Logistic Regression.

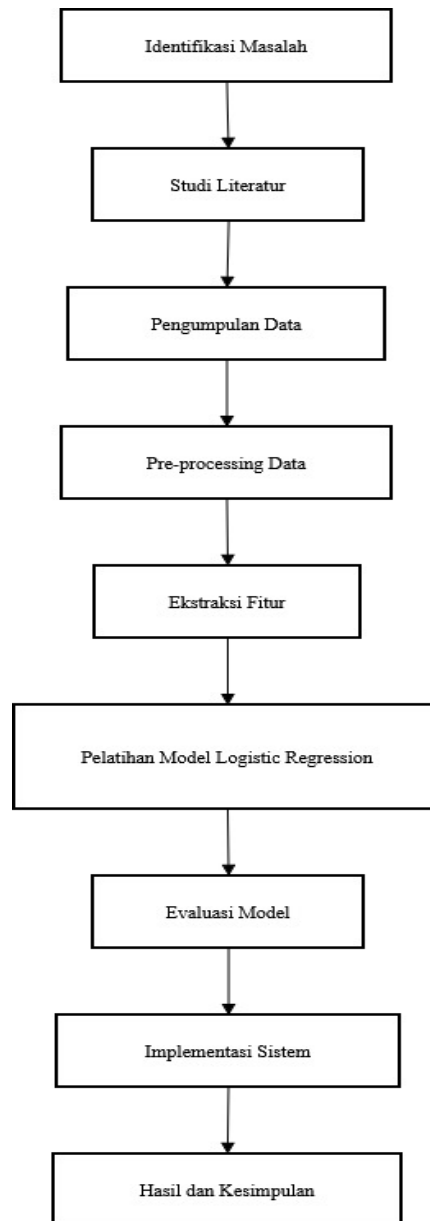
Model Logistic Regression pada penelitian ini digunakan untuk permasalahan klasifikasi biner, yaitu membedakan tingkat kematangan alpukat ke dalam dua kelas, yaitu belum matang dan matang.

Meskipun pada dataset awal terdapat lima label kematangan (hard, pre-conditioned, breaking, firm-ripe, dan ripe), seluruh label tersebut kemudian digabungkan menjadi dua kelas utama. Label hard, pre-conditioned, dan breaking dikategorikan sebagai belum matang, sedangkan label firm-ripe dan ripe dikategorikan sebagai matang.

Transformasi ini dilakukan untuk menyesuaikan kebutuhan klasifikasi biner serta meningkatkan kestabilan model Logistic Regression.

A. Tahapan Penelitian

Penelitian ini dilakukan melalui serangkaian tahapan yang terstruktur, dimulai dari pengumpulan dataset karakteristik buah alpukat, dilanjutkan dengan proses pra-pemrosesan data, normalisasi dan transformasi fitur, serta pelatihan dan pengujian model prediksi menggunakan algoritma Logistic Regression. Seluruh tahapan tersebut disusun secara sistematis untuk memastikan proses pemodelan berjalan dengan baik dan menghasilkan prediksi tingkat kematangan buah alpukat yang objektif dan konsisten.



Gambar 1. Alur Penelitian

B. Dataset Penelitian

Dataset yang digunakan dalam penelitian ini berupa data numerik karakteristik fisik buah alpukat, bukan citra digital. Setiap data merepresentasikan satu sampel buah alpukat yang diukur berdasarkan beberapa parameter fisik, yaitu firmness (kekerasan), hue, saturation, brightness, sound (dB), berat (gram), dan ukuran (cm³).

Dataset ini digunakan sebagai input untuk memprediksi tingkat kematangan alpukat ke dalam dua kelas, yaitu belum matang dan matang, menggunakan algoritma Logistic Regression.

Tabel 1. Dataset Alpukat

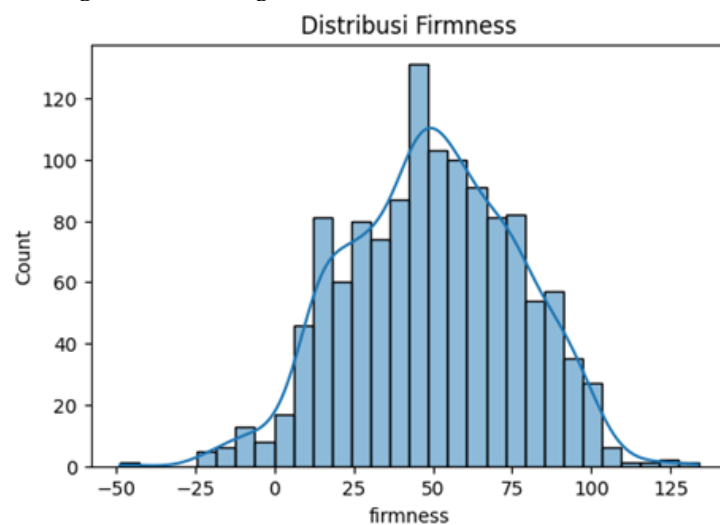
No	Firmness	Hue	Saturation	Brightness	Color_category	Sound_db	Wight_g	Size_cm3	Ripeness
1	14.5	19	40	26	Black	34	175	261	Ripe
2	71.7	53	69	75	Green	69	206	185	Pre-conditioned
3	88.5	60	94	46	Dark Green	79	220	143	Hard
4	27	312	52	30	Purple	51	169	224	Firm-ripe
5	42.5	303	58	32	Purple	63	200	227	Breaking

Tabel 1 menyajikan contoh data karakteristik buah alpukat beserta label tingkat kematangan awal. Label Ripeness pada tabel tersebut merupakan label asli sebelum dilakukan proses transformasi. Pada tahap preprocessing, label-label tersebut kemudian dipetakan ke dalam dua kelas utama, yaitu belum matang dan matang, sesuai dengan kebutuhan pemodelan klasifikasi biner menggunakan algoritma Logistic Regression.

C. Statistik Deskriptif Dataset

Analisis statistik deskriptif dilakukan untuk memberikan gambaran awal mengenai pola sebaran dan karakteristik atribut numerik pada dataset yang digunakan. Pada penelitian ini, visualisasi distribusi difokuskan pada atribut firmness, hue, dan sound_db, karena ketiga atribut tersebut berperan penting dalam merepresentasikan karakteristik fisik alpukat yang berkaitan langsung dengan tingkat kematangan. Histogram beserta kurva distribusi pada masing-masing atribut digunakan sebagai dasar analisis lebih lanjut.

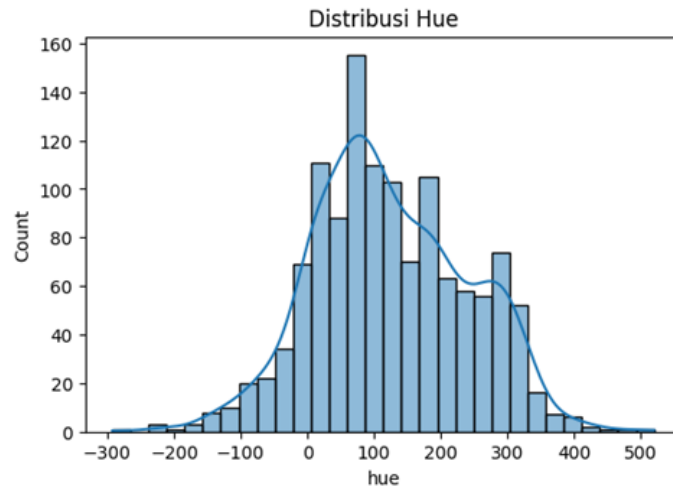
1. Distribusi firmness menunjukkan bahwa nilai kekerasan alpukat tersebar cukup luas dengan konsentrasi utama berada pada rentang nilai menengah. Pola histogram yang cenderung mendekati distribusi normal mengindikasikan bahwa sebagian besar sampel memiliki tingkat kekerasan yang berada di sekitar nilai rata-rata, sementara nilai ekstrem relatif lebih sedikit. Sebaran ini menunjukkan bahwa dataset mencakup variasi kondisi alpukat dari yang masih keras hingga yang lebih lunak, sehingga fitur firmness memiliki kontribusi penting dalam merepresentasikan perbedaan tingkat kematangan buah.



Gambar 2. Distribusi Frimness

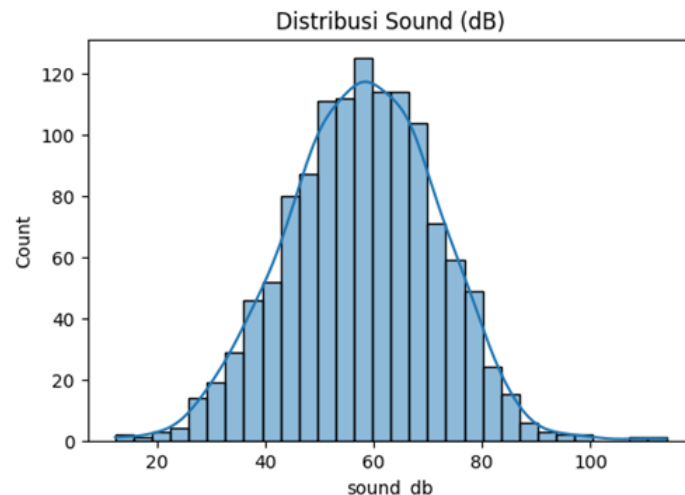
2. Distribusi nilai hue memperlihatkan penyebaran data yang cukup lebar dengan beberapa puncak pada rentang tertentu, yang mencerminkan variasi warna kulit alpukat dalam dataset. Variasi ini menggambarkan perubahan warna yang terjadi

selama proses pematangan, mulai dari warna hijau cerah hingga warna yang lebih gelap. Pola distribusi yang tidak sepenuhnya simetris menunjukkan adanya keragaman kondisi visual alpukat, sehingga fitur hue dapat memberikan informasi yang signifikan bagi model dalam membedakan tingkat kematangan berdasarkan karakteristik warna.



Gambar 3. Distribusi Hue

3. Distribusi nilai `sound_db` menunjukkan pola yang relatif simetris dan mendekati distribusi normal, dengan sebagian besar data terkonsentrasi di sekitar nilai tengah. Hal ini mengindikasikan bahwa mayoritas sampel alpukat menghasilkan intensitas suara ketukan yang relatif seragam, sementara variasi yang muncul mencerminkan perbedaan struktur internal dan kepadatan daging buah. Variasi pada nilai `sound_db` ini relevan untuk digunakan sebagai salah satu indikator kematangan, karena perubahan karakteristik internal buah dapat memengaruhi respons suara yang dihasilkan saat dilakukan pengetukan.



Gambar 4. Distribusi Sound (dB)

D. Preprocessing Data

Pada tahap preprocessing, dilakukan penyesuaian terhadap variabel target agar sesuai dengan kebutuhan pemodelan. Pada dataset awal, variabel *ripeness* memiliki lima kelas kematangan, yaitu *hard*, *pre-conditioned*, *breaking*, *firm-ripe*, dan *ripe*. Struktur multikelas ini meningkatkan kompleksitas klasifikasi dan berpotensi memengaruhi kestabilan model Logistic Regression, khususnya ketika jumlah data terbatas.

Tabel 2. Perubahan Struktur Label Dataset Alpukat

Aspek	Data Lama	Data Baru
Bentuk Label	Multikelas	Biner
Jumlah Kelas	5 Kelas	2 Kelas
Kompleksitas Model	Tinggi	Lebih Rendah
Risiko Overfitting	Lebih Besar	Lebih Kecil
Kesesuaian Logistic Regression	Cukup	Lebih Optimal

Untuk menyederhanakan permasalahan, variabel target kemudian ditransformasikan menjadi klasifikasi biner dengan membentuk variabel *ripeness_binary* yang terdiri dari dua kelas, yaitu *belum_matang* dan *matang*. Transformasi ini bertujuan untuk meningkatkan fokus model dalam membedakan alpukat yang siap dan belum siap dikonsumsi, sekaligus menurunkan risiko overfitting dan meningkatkan kesesuaian penerapan Logistic Regression.

```

▶ print(df['ripeness_binary'].value_counts())
print("Jumlah NaN:", df['ripeness_binary'].isna().sum())

... ripeness_binary
    belum_matang    1005
      matang         245
Name: count, dtype: int64
Jumlah NaN: 0

```

Gambar 5. Ripeness Binary

Hasil preprocessing menunjukkan bahwa dataset terdiri dari 1005 data berlabel *belum_matang* dan 245 data berlabel *matang*, serta tidak ditemukan nilai kosong (NaN) pada variabel target. Berdasarkan Gambar 4.4, terlihat bahwa distribusi kelas tidak seimbang, dengan dominasi kelas *belum_matang*. Meskipun demikian, struktur data hasil preprocessing dinilai lebih sederhana, stabil, dan siap digunakan pada tahap pelatihan model Logistic Regression berbasis klasifikasi biner.

E. Ekstraksi Fitur

Tahap ekstraksi fitur dilakukan setelah proses *pre-processing* dengan tujuan membentuk representasi numerik akhir dari setiap sampel alpukat yang akan digunakan sebagai masukan model klasifikasi. Pada penelitian ini, fitur diekstraksi dari dua sumber utama, yaitu data pengukuran fisik alpukat dan representasi numerik karakteristik warna kulit buah.

Tabel 3. Vektor Fitur Ekstraksi

No	Nama Fitur	Jenis Data	Nilai Contoh	Keterangan
1	Firmness	Numerik	0.42	Nilai kekerasan kulit alpukat setelah standardisasi
2	Hue	Numerik	-0.31	Nilai hue hasil ekstraksi warna yang telah distandarisasi
3	Saturation	Numerik	0.15	Tingkat kejenuhan warna kulit alpukat
4	Brightness	Numerik	0.28	Tingkat kecerahan warna kulit alpukat

5	Sound_db	Numerik	0.66	Intensitas suara hasil ketukan alpukat
6	Weight_g	Numerik	-0.12	Berat alpukat setelah proses standardisasi
7	Size_cm3	Numerik	0.08	Ukuran atau volume alpukat setelah standardisasi
8	Color_green	Biner	1	Hasil one-hot encoding warna hijau
9	Color_dark_green	Biner	0	Hasil one-hot encoding warna hijau tua

Tabel 3. menunjukkan contoh vektor fitur akhir yang dihasilkan setelah proses ekstraksi dan transformasi data. Seluruh fitur numerik telah melalui proses standardisasi sehingga memiliki skala yang seragam, sedangkan fitur kategorikal direpresentasikan dalam bentuk variabel biner melalui teknik *one-hot encoding*. Vektor fitur ini digunakan sebagai masukan pada algoritma Logistic Regression dalam proses pelatihan dan prediksi tingkat kematangan alpukat.

HASIL DAN PEMBAHASAN

Bagian ini menyajikan hasil pengujian sistem prediksi tingkat kematangan buah alpukat yang telah dikembangkan serta pembahasan kinerja model berdasarkan metrik evaluasi yang digunakan. Evaluasi dilakukan untuk menilai tingkat akurasi dan konsistensi model dalam mengklasifikasikan tingkat kematangan alpukat. Sebelum dilakukan analisis hasil klasifikasi, model divalidasi menggunakan pembagian data latih dan data uji dengan rasio 75:25, sehingga hasil pengujian dapat merepresentasikan kemampuan generalisasi model secara objektif.

A. Pembagian Data

Dataset alpukat yang digunakan dalam penelitian ini dibagi menjadi dua bagian utama, yaitu data latih dan data uji. Pembagian data bertujuan untuk memastikan bahwa proses pelatihan dan evaluasi model dilakukan secara terpisah, sehingga kinerja model dapat diukur secara objektif pada data yang belum pernah digunakan sebelumnya. Proporsi pembagian data mengikuti praktik umum dalam penelitian machine learning, yaitu sekitar 75% untuk data latih dan 25% untuk data uji.

Tabel 4 . Pembagian Data

Jenis Data	Jumlah Data	Persentase
Data Latih	937	74.96%
Data Uji	313	25.04%
Total	1250	100%

Data latih digunakan pada proses pembentukan model prediksi tingkat kematangan alpukat menggunakan algoritma Logistic Regression, sedangkan data uji digunakan untuk mengevaluasi kinerja model. Pembagian ini bertujuan untuk mengurangi risiko overfitting serta memberikan gambaran kemampuan model dalam melakukan prediksi pada data baru.

B. Hasil Klasifikasi

Hasil pengujian menunjukkan bahwa sistem prediksi tingkat kematangan alpukat mampu mencapai tingkat akurasi sekitar 77%. Hasil ini menunjukkan bahwa algoritma Logistic Regression dengan fitur numerik yang digunakan, seperti firmness, hue, saturation, brightness,

sound_db, berat, dan ukuran buah, mampu memberikan kinerja yang cukup baik dalam membedakan alpukat matang dan belum matang.

Selain akurasi, evaluasi kinerja model juga dilakukan menggunakan *classification report* yang mencakup metrik precision, recall, dan f1-score untuk masing-masing kelas. Evaluasi ini bertujuan untuk memberikan gambaran yang lebih rinci mengenai kemampuan model dalam mengklasifikasikan setiap tingkat kematangan alpukat.

...	precision	recall	f1-score	support
0	0.84	0.88	0.86	252
1	0.38	0.30	0.33	61
accuracy			0.77	313
macro avg	0.61	0.59	0.60	313
weighted avg	0.75	0.77	0.76	313
Akurasi: 0.7699680511182109				

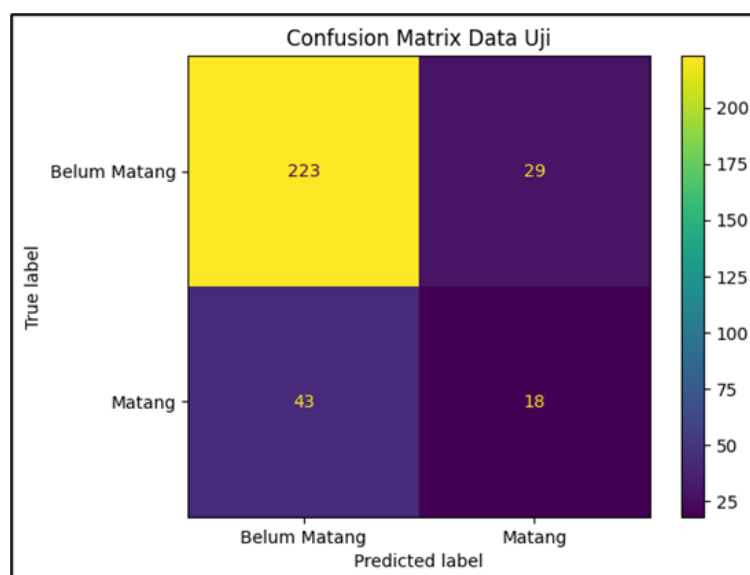
Gambar 6. Classification Report

Berdasarkan hasil *classification report*, model menunjukkan performa yang relatif seimbang pada kedua kelas. Kelas *belum_matang* memiliki nilai precision dan recall yang lebih tinggi, yang dipengaruhi oleh jumlah data yang lebih dominan pada kelas tersebut. Sementara itu, kelas *matang* menunjukkan nilai recall yang sedikit lebih rendah, yang mengindikasikan adanya beberapa data alpukat matang yang masih terklasifikasi sebagai belum matang. Meskipun demikian, nilai *macro average* dan *weighted average* yang mendekati nilai akurasi keseluruhan menunjukkan bahwa model memiliki performa klasifikasi yang cukup konsisten.

Secara keseluruhan, hasil ini menunjukkan bahwa Logistic Regression dapat digunakan sebagai pendekatan yang sederhana, efisien, dan cukup andal dalam sistem prediksi tingkat kematangan alpukat berbasis data numerik, meskipun masih terdapat ruang untuk peningkatan performa melalui penambahan data dan pengembangan metode lebih lanjut.

C. Confusion Matrix

Untuk mengevaluasi performa model dalam mengklasifikasikan tingkat kematangan buah alpukat, dilakukan analisis menggunakan *confusion matrix*. *Confusion matrix* digunakan untuk menampilkan jumlah prediksi yang benar dan salah pada masing-masing kelas, yaitu *belum_matang* dan *matang*, sehingga pola kesalahan klasifikasi dapat dianalisis secara lebih rinci.



Gambar 7. Confusion Matrix

Berdasarkan *confusion matrix* pada Gambar 7, model Logistic Regression mampu mengklasifikasikan sebagian besar data alpukat dengan benar. Pada kelas *belum_matang*, sebanyak 223 data diprediksi dengan tepat, sementara 29 data salah diklasifikasikan sebagai *matang*. Pada kelas *matang*, terdapat 18 data yang berhasil dikenali dengan benar dan 43 data yang masih diprediksi sebagai *belum_matang*. Kesalahan ini menunjukkan bahwa model lebih sulit membedakan alpukat matang, yang dipengaruhi oleh kemiripan karakteristik fisik, khususnya pada fitur firmness dan sound_db. Meskipun demikian, jumlah prediksi benar lebih dominan dibandingkan prediksi salah, sejalan dengan akurasi model sekitar 77%.

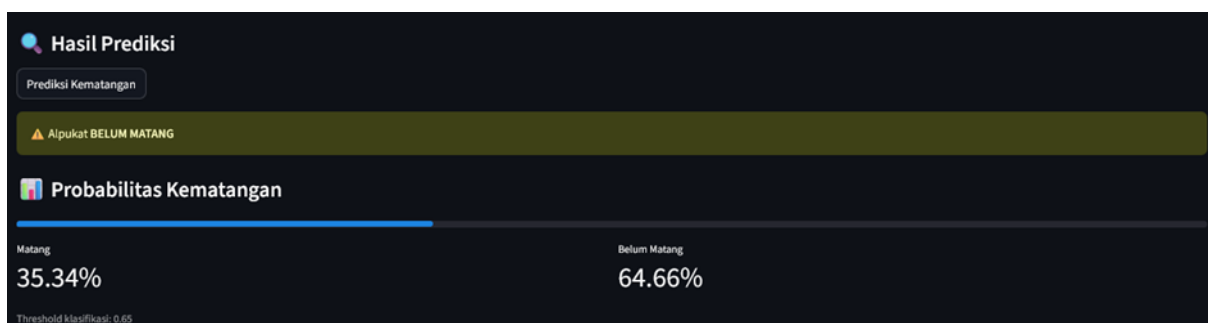
D. Implementasi Sistem

Implementasi sistem prediksi tingkat kematangan buah alpukat dilakukan melalui aplikasi web yang dikembangkan menggunakan framework Streamlit. Aplikasi ini dirancang agar dapat digunakan secara langsung oleh pengguna tanpa memerlukan instalasi perangkat lunak tambahan. Antarmuka web memungkinkan pengguna melakukan prediksi tingkat kematangan alpukat secara sederhana dan interaktif.

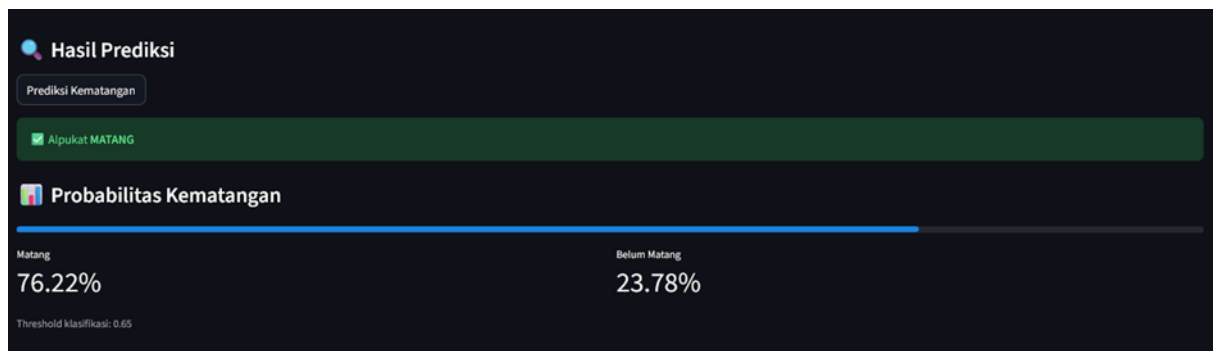


Gambar 8. Antarmuka Input Prediksi Kematangan Buah Alpukat

Pengguna dapat memasukkan data karakteristik fisik buah alpukat melalui halaman aplikasi. Data tersebut kemudian diproses secara otomatis melalui tahapan preprocessing dan standardisasi fitur numerik, sebelum dilakukan klasifikasi menggunakan model Logistic Regression yang telah dilatih. Hasil prediksi ditampilkan dalam bentuk status kematangan alpukat, yaitu *belum matang* atau *matang*, beserta nilai probabilitas sebagai tingkat kepercayaan model.



Gambar 9. Hasil Prediksi Kematangan Buah Alpukat Belum Matang



Gambar 10. Hasil Prediksi Kematangan Buah Alpukat Matang

Selain itu, sistem juga dilengkapi dengan mekanisme validasi untuk memastikan bahwa data yang dimasukkan pengguna berada dalam rentang nilai yang sesuai dengan dataset pelatihan. Mekanisme ini bertujuan untuk mengurangi kesalahan prediksi akibat input data yang tidak realistis, sehingga sistem dapat memberikan hasil prediksi yang lebih andal.

KESIMPULAN

Penelitian ini berhasil menghasilkan sistem prediksi tingkat kematangan buah alpukat berbasis data numerik dengan memanfaatkan algoritma Logistic Regression. Sistem yang dikembangkan mampu mengklasifikasikan tingkat kematangan alpukat ke dalam dua kelas, yaitu *belum matang* dan *matang*, serta diimplementasikan dalam bentuk aplikasi web berbasis Streamlit yang mudah digunakan oleh pengguna.

Hasil pengujian menunjukkan bahwa sistem mencapai tingkat akurasi sekitar 77%, dengan nilai precision, recall, dan f1-score yang cukup baik pada kedua kelas kematangan. Selain mampu memprediksi tingkat kematangan alpukat, sistem juga menampilkan nilai probabilitas prediksi sebagai tingkat kepercayaan model, sehingga hasil yang diberikan lebih informatif dan dapat mendukung pengambilan keputusan.

Meskipun menunjukkan kinerja yang cukup baik, sistem masih memiliki keterbatasan, terutama pada kemiripan karakteristik fisik antara alpukat yang hampir matang dan yang belum matang, serta keterbatasan jumlah dan variasi dataset. Oleh karena itu, pengembangan selanjutnya dapat difokuskan pada penambahan data, peningkatan variasi fitur, serta penerapan metode klasifikasi lain atau pendekatan hybrid untuk meningkatkan akurasi dan kemampuan generalisasi sistem.

DAFTAR PUSTAKA

- [1] I. Sutoyo, "IMPLEMENTASI ALGORITMA DECISION TREE UNTUK KLASIFIKASI DATA PESERTA DIDIK," vol. 14, no. 2, 2018, [Online]. Available: www.bsi.ac.id
- [2] H. Annur, "KLASIFIKASI MASYARAKAT MISKIN MENGGUNAKAN METODE NAÏVE BAYES," 2018.
- [3] A. Samosir, M. Hasibuan, W. E. Justino, and T. Hariyono, "Komparasi Algoritma Random Forest, Naïve Bayes dan K-Nearest Neighbor Dalam klasifikasi Data Penyakit Jantung".
- [4] D. Armiaedy and I. M. R, "Klasifikasi Kualitas Buah Pisang Berdasarkan Citra Buah Menggunakan Stochastic Gradient Descent," *KLIK: Kajian Ilmiah Informatika dan Komputer*, vol. 4, no. 2, 2023.
- [5] W. F. Hidayat, T. Asra, and A. Setiadi, "Klasifikasi Penyakit Daun Kentang Menggunakan Model Logistic Regression," *Indonesian Journal on Software Engineering (IJSE)*, vol. 8, no. 2, pp. 173–179, 2022, [Online]. Available: <http://ejournal.bsi.ac.id/ejurnal/index.php/ijse>
- [6] H. Hikmayanti Handayani, K. Ahmad Baihaqi, and U. Buana Perjuangan Karawang, "Implementasi Algoritma Logistic Regression Untuk Klasifikasi Penyakit Stroke," 2023.

- [7] D. Najwa Ardelia, H. Desfianty Arifin, S. Daniswara, A. Puspita Sari, P. Studi Informatika -UPN, and J. Timur Jl Raya Rungkut Madya Gunung Anyar -Surabaya, "Klasifikasi Harga Ponsel Menggunakan Algoritma Logistic Regression."
- [8] A. Adrian, I. Verawati,) Program, and S. Informatika, "Analisis Performa Logistic Regression dan Random Forest dalam Klasifikasi Kelayakan Penerimaan Kredit," 2025. [Online]. Available: <https://subset.id/index.php/IJCSR>
- [9] I. R. Muslem and T. M. Johan, "KLIK: Kajian Ilmiah Informatika dan Komputer Klasifikasi Citra Ikan Menggunakan Algoritma Convolutional Neural Network dengan Arsitektur VGG-16," *Media Online*, vol. 4, no. 2, pp. 978–985, 2023, doi: 10.30865/klik.v4i2.1209.
- [10] A. Hamzah, E. Susanti, and R. M. Lestari, "SUPPORT VECTOR MACHINE (SVM)".
- [11] M. Noer, F. Hidayat, J. K. Zaini Mun'im, P. N. Jadid, and J. Timur, "KLASIFIKASI BUAH ALPUKAT BERDASARKAN TEKSTUR BUAH MENGGUNAKAN METODE BACKPROPAGATION BERBASIS IMAGE PROCESSING," 2023. [Online]. Available: <http://e-journal.stmiklombok.ac.id/index.php/jirelISSN.2620-6900>
- [12] A. Ayaz Mirani, M. Suleman Memon, R. Chohan, A. Ali Wagan, and M. Qabulio, "Machine Learning In Agriculture: A Review," *LUME*, vol. 10, p. 5, 2021, [Online]. Available: www.ijstr.org
- [13] C. H. H. Jannah, I. Muslem, and D. Azmi, "Jurnal Ilmu Komputer Aceh Klasifikasi Plat Nomor Kendaraan Berdasarkan Wilayah Tertentu Menggunakan Algoritma Optical Character Recognition," *Jurnal Ilmu Komputer Aceh*, Oct. 2025, [Online]. Available: <https://jurnal.fikompublisher.com/ilka/article/view/16>
- [14] Gareth. James, Daniela. Witten, Trevor. Hastie, and Robert. Tibshirani, *An introduction to statistical learning : with applications in R*. Springer : Springer Science+Business Media, 2017.
- [15] P. O. Akinwumi, S. Ojo, T. I. Nathaniel, J. Wanliss, O. Karunwi, and M. Sulaiman, "Evaluating machine learning models for stroke prediction based on clinical variables," *Front. Neurol.*, vol. 16, 2025, doi: 10.3389/fneur.2025.1668420.
- [16] I. H. Ikasari, P. Rosyani, and R. Amalia, "Klasifikasi Jenis Buah Menggunakan Metode CNN," *RIGGS: Journal of Artificial Intelligence and Digital Business*, vol. 4, no. 2, pp. 5451–5458, Jul. 2025, doi: 10.31004/riggs.v4i2.1271.
- [17] G. Honestya, M. Sajida, and A. Ramadhanu, "Klasifikasi Jenis Daun Herbal Menggunakan Metode Logistic Regression dan Decision Tree Classifier Berdasarkan Fitur (Warna dan Bentuk)".
- [18] F. B. Setiawan, C. B. Adipradana, and L. H. Pratomo, "Fruit Ripeness Classification System Using Convolutional Neural Network (CNN) Method," *PROtek : Jurnal Ilmiah Teknik Elektro*, vol. 10, no. 1, p. 46, Jan. 2023, doi: 10.33387/protk.v10i1.5549.
- [19] M. Knott, F. Perez-Cruz, and T. Defraeye, "Facilitated machine learning for image-based fruit quality assessment," *J. Food Eng.*, vol. 345, May 2023, doi: 10.1016/j.jfoodeng.2022.111401.
- [20] A. Hamzah, E. Susanti, and R. M. Lestari, "SUPPORT VECTOR MACHINE (SVM)".
- [21] A. Kamilaris and F. X. Prenafeta-Boldú, "Deep Learning in Agriculture: A Survey."
- [22] L. Luthfi, R. Imam Muslem, D. Armady, S. Sriwinar, R. Fajri, and I. Iqbal, "Analysis of CNN Method for Image Classification of Coconut Ripeness Levels," in *2023 Eighth International Conference on Informatics and Computing (ICIC)*, IEEE, Dec. 2023, pp. 1–6. doi: 10.1109/ICIC60109.2023.10381964.
- [23] A. Caban, K. K. April Gregorio, K. V. Macam, M. S. Rosario Puzon, C. S. Santillan, and M. C. Santillan, "CLASSIFICATION OF NIPA FRUIT USING ARTIFICIAL NEURAL NETWORK," *Asian Journal of Multidisciplinary Studies*, vol. 5, no. 1, 2022.
- [24] D. M. Bulanon and T. Kataoka, "Fruit detection system and an end effector for robotic harvesting of Fuji apples," 2010. [Online]. Available: <http://www.cigrjournal.org>

- [25] N. Utami Putri and E. Redi Susanto, "Klasifikasi Jenis Kayu Menggunakan Support Vector Machine Berdasarkan Ciri Tekstur Local Binary Pattern," *CYBERNETICS*, vol. 4, no. 02, pp. 93–100, 2020.
- [26] J. Liu, J. Sun, Y. Wang, X. Liu, Y. Zhang, and H. Fu, "Non-Destructive Detection of Fruit Quality: Technologies, Applications and Prospects," Jun. 01, 2025, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi: 10.3390/foods14122137.
- [27] F. D. Pramakrisna, F. D. Adhinata, and N. A. F. Tanjung, "Aplikasi Klasifikasi SMS Berbasis Web Menggunakan Algoritma Logistic Regression," *Teknika*, vol. 11, no. 2, pp. 90–97, Jun. 2022, doi: 10.34148/teknika.v11i2.466.
- [28] B. Tri Putra, E. Yulianingsih, S. Informasi, S. Teknologi, and U. Bina Darma, "Analisis Tingkat Akurasi Prediksi Gejala COVID-19 Dengan Menggunakan Metode Logistic Regression dan Support Vector Machine", [Online]. Available: <http://github.com/nshomron/covidpred>.
- [29] W. T. Setiadi, D. Jollyta, and E. B. Setiawan, "Seminar Nasional Informatika (SENATIKA) Prosiding Senatika 2025 Prediksi Risiko Diabetes Menggunakan Model Regresi Logistik dan Random Forest."