INCODING: Journal of Informatic and Computer Science Engineering
https://journal.mahesacenter.org/index.php/incoding/index || ISSN 2776-432X (online)
6(1) 2026: 35-49

DOI: https://doi.org/10.34007/incoding.v6i1.930



Penerapan Algoritma K-Nearest Neighbor dalam Menentukan Kualitas Buah Jeruk Bali

Application of K – Nearest Neighbor Algorithm in Determining the Quality of Bali Pomelo Fruit

Musbar Fandi Harahap, Fera Damayanti, & Siti Sundari
Teknik Informatika, Teknik dan Komputer, Universitas Harapan Medan, Indonesia

Diterima: 2025-06-29; Direview: 2026-04-30; Disetujui: 2026-06-04

*Coresponding Email: musbarfandi2123@gmail.com

Abstrak

Jeruk bali (Citrus maxima) merupakan tanaman buah tropis bernilai ekonomi tinggi dan berpotensi dikembangkan di Indonesia. Namun, penentuan kualitas buah masih banyak dilakukan secara manual sehingga hasil penilaian cenderung subjektif. Penelitian ini bertujuan menerapkan algoritma K-Nearest Neighbor (K-NN) untuk menentukan kualitas jeruk bali berdasarkan atribut fisik buah. Variabel yang digunakan meliputi pigmen kulit buah, kemulusan, kelunakan buah, bobot buah, dan tebal kulit buah, sedangkan label hasil diklasifikasikan menjadi kategori baik dan tidak baik. Data penelitian terdiri atas 50 data latih dan 2 data prediksi. Tahapan penelitian meliputi analisis data, preprocessing, pembagian data uji, perhitungan jarak menggunakan Euclidean Distance, serta pengujian menggunakan aplikasi Orange. Hasil pengujian menunjukkan dua data prediksi diklasifikasikan sebagai buah berkualitas baik. Evaluasi model memperoleh nilai akurasi 0,834, classification accuracy 0,860, F1 0,860, precision 0,861, recall 0,860, dan MCC 0,720.
Kata Kunci: Jeruk Bali; Klasifikasi; K-Nearest Neighbor; Data Mining; Orange.

Abstract

Pomelo (Citrus maxima) is a tropical fruit crop with high economic value and strong potential for development in Indonesia. However, fruit quality assessment is still commonly conducted manually, which may lead to subjective results. This study aims to apply the K-Nearest Neighbor (K-NN) algorithm to determine pomelo fruit quality based on its physical attributes. The variables used include fruit skin pigment, smoothness, fruit softness, fruit weight, and peel thickness, while the output label is classified into good and poor quality categories. The dataset consists of 50 training data and 2 prediction data. The research stages include data analysis, preprocessing, test data splitting, distance calculation using Euclidean Distance, and testing using the Orange application. The results show that the two prediction data were classified as good-quality fruit. The model evaluation obtained an accuracy of 0.834, classification accuracy of 0.860, F1 score of 0.860, precision of 0.861, recall of 0.860, and MCC of 0.720.
Keywords: *Pomelo; Classification; K-Nearest Neighbor; Data Mining; Orange.*

How To Cite: Harahap, M. F., Damayanti, F., & Sundari, S. (2026). Penerapan Algoritma K-Nearest Neighbor dalam Menentukan Kualitas Buah Jeruk Bali. INCODING: Journal of Informatic and Computer Science Engineering, 6(1) 2026: 35-49

This work is licensed under a Creative Commons Attribution 4.0

PENDAHULUAN

Jeruk bali, jeruk besar, atau pomelo (Citrus maxima) merupakan salah satu tanaman buah tropis yang memiliki nilai ekonomi tinggi dan berpotensi untuk terus dikembangkan di Indonesia [1]. Tanaman ini tersebar di berbagai wilayah dan memiliki beragam kelompok jenis, antara lain Bali Merah, Cikoneng, Nambangan, Raja, Ratu, dan Pangkep. Jeruk bali juga dikenal sebagai salah satu jenis jeruk lokal penting selain Citrus medica dan Citrus reticulata. Sebagai tanaman yang tumbuh baik di daerah tropis, jeruk bali memiliki peluang pengembangan yang cukup besar, baik dari sisi budidaya, produktivitas, maupun peningkatan kualitas hasil panen [2], [3].(Fauziah, 2022)(MARHAWATI, 2019; Pijpe et al., 2025)

Secara fisik, jeruk bali memiliki ciri khas berupa ukuran buah yang relatif besar, rasa yang segar, serta daya simpan yang cukup lama dalam kondisi tertentu [4]. Selain itu, jeruk bali juga memiliki kandungan gizi yang bermanfaat bagi kesehatan. Buah ini mengandung senyawa antioksidan, seperti fenol dan flavonoid, serta vitamin C dan vitamin A yang berperan bagi kesehatan. Beberapa karakteristik buah jeruk bali yang dianggap unggul antara lain warna daging buah yang menarik, rasa getir yang rendah, mudah dikupas, serta jumlah biji yang sedikit atau bahkan tidak berbiji [5].(Makrufa, 2025)(Bagaskara, 2021)

Meskipun memiliki potensi ekonomi dan agronomis yang tinggi, pengembangan jeruk bali di Indonesia masih menghadapi beberapa kendala, terutama dalam hal penentuan kualitas buah secara objektif dan konsisten. Kualitas buah jeruk bali umumnya dapat diamati melalui beberapa karakteristik, seperti ukuran dan bentuk buah, bentuk ujung dan pangkal buah, warna dan tekstur flavedo atau kulit luar buah, ketebalan serta warna albedo, warna endokarpium, warna dan rasa vesikula atau daging buah, aroma minyak atsiri, jumlah buah per pohon, serta jumlah biji pada setiap buah [6]. Karakteristik tersebut menjadi dasar penting dalam proses seleksi untuk memperoleh buah dengan kualitas yang baik.(Rai et al., 2026)

Dalam praktiknya, penentuan kualitas buah jeruk bali masih sering dilakukan secara manual. Proses tersebut biasanya dilakukan dengan mengamati kondisi fisik buah secara langsung, melihat permukaan kulit buah, serta memukul buah secara perlahan untuk memperkirakan tingkat kematangan. Setelah itu, data buah yang diperoleh dimasukkan ke dalam aplikasi sederhana. Namun, data yang telah dikumpulkan tersebut belum dimanfaatkan secara optimal untuk proses analisis. Akibatnya, penentuan kualitas buah

This work is licensed under a Creative Commons Attribution 4.0

masih sangat bergantung pada pengalaman subjektif, sehingga berpotensi menimbulkan perbedaan penilaian antara satu pengamat dengan pengamat lainnya.

Permasalahan tersebut menunjukkan perlunya pendekatan yang lebih sistematis dalam menentukan kualitas buah jeruk bali. Salah satu pendekatan yang dapat digunakan adalah data mining. Data mining merupakan proses penggalian informasi dari kumpulan data untuk menemukan pola, hubungan, atau pengetahuan yang dapat digunakan dalam proses pengambilan keputusan [7], [8]. Dalam konteks penentuan kualitas buah, data mining dapat membantu mengolah data karakteristik buah agar menghasilkan klasifikasi kualitas yang lebih terukur dan konsisten.(Jollyta et al., 2020; Khormarudin, 2016)

Salah satu teknik data mining yang dapat digunakan untuk proses klasifikasi adalah algoritma K-Nearest Neighbor (K-NN). Algoritma K-NN mengklasifikasikan data baru berdasarkan kedekatannya dengan data latih yang sudah memiliki label kelas. Data baru dibandingkan dengan data yang tersedia melalui perhitungan jarak, kemudian kelasnya ditentukan berdasarkan mayoritas kelas dari sejumlah tetangga terdekat [9], [10]. Metode ini relatif sederhana dan mudah diterapkan. Oleh karena itu, algoritma ini dapat dipertimbangkan untuk klasifikasi kualitas buah jeruk bali berdasarkan variabel yang telah ditentukan [11].(Bakri & Harahap, 2025; Fahmi Ruziq et al., 2026)(FERI, 2025)

Namun demikian, algoritma K-NN juga memiliki kelemahan, terutama dari sisi komputasi. Proses klasifikasi membutuhkan perhitungan jarak antara data uji dengan data latih, sehingga waktu komputasi dapat meningkat apabila jumlah data semakin besar [12], [13]. Meskipun demikian, dalam penelitian ini algoritma K-NN tetap relevan digunakan karena jumlah data yang dianalisis masih dapat dikelola dan variabel yang digunakan dapat direpresentasikan dalam bentuk numerik.(Ujianto et al., 2025; Wardhani et al., 2025)

Beberapa penelitian terdahulu menunjukkan bahwa algoritma K-NN telah digunakan dalam proses klasifikasi kualitas buah maupun identifikasi tingkat kematangan [14], [15]. Pawening et al. menerapkan K-NN untuk klasifikasi kualitas jeruk lokal berdasarkan tekstur dan bentuk, dengan tingkat akurasi tertinggi sebesar 93,33% dan terendah sebesar 86,20% [16]. Penelitian lain membandingkan metode Convolutional Neural Network (CNN) dan K-NN dalam klasifikasi buah jeruk, dengan hasil bahwa K-NN menghasilkan akurasi lebih rendah dibandingkan CNN [17]. Selain itu, penelitian tentang identifikasi kematangan buah jeruk Medan menggunakan K-NN berbasis fitur Red, Green,

This work is licensed under a Creative Commons Attribution 4.0

Blue (RGB) melaporkan akurasi sistem sebesar 96,6% [18].(Guslianto, 2023; Sanu, 2025)Pawening et al. (2020)Siwilopo & Marcos (2023)SURYATAMA JAYA PUTRA (2023)

Penelitian ini dilakukan untuk menerapkan algoritma K-Nearest Neighbor dalam menentukan kualitas buah jeruk bali. Penelitian ini menggunakan sejumlah variabel yang berkaitan dengan karakteristik buah, yaitu pigmen kulit buah, kemulusan, kelunakan buah, bobot buah, dan tebal kulit buah. Melalui penerapan algoritma K-NN, diharapkan proses penentuan kualitas buah jeruk bali dapat dilakukan secara lebih sistematis, objektif, dan terukur, sehingga dapat membantu proses seleksi kualitas buah berdasarkan data yang tersedia.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode klasifikasi data mining. Objek yang dianalisis adalah data kualitas buah jeruk bali berdasarkan beberapa variabel fisik buah, yaitu pigmen kulit buah, kemulusan, kelunakan buah, bobot buah, dan tebal kulit buah. Variabel-variabel tersebut digunakan sebagai atribut dalam proses klasifikasi, sedangkan hasil kualitas buah digunakan sebagai label kelas, yaitu “baik” dan “tidak baik”. Data yang digunakan dalam penelitian ini berjumlah 50 data latih dan 2 data prediksi yang digunakan untuk menguji proses klasifikasi kualitas buah jeruk bali.

Tahapan penelitian diawali dengan pengumpulan data karakteristik buah jeruk bali. Data yang diperoleh kemudian dianalisis untuk mengetahui struktur data, jumlah record, jenis atribut, serta label kelas yang digunakan. Setelah itu, dilakukan proses konversi terhadap data kategorik agar dapat dibaca oleh sistem. Label kualitas buah “baik” dikodekan menjadi angka 1, sedangkan “tidak baik” dikodekan menjadi angka 2. Proses ini dilakukan karena algoritma klasifikasi membutuhkan data dalam bentuk numerik agar dapat dihitung dan dianalisis secara komputasional.

Sebelum data diproses menggunakan algoritma K-Nearest Neighbor, dilakukan tahapan preprocessing. Tahapan ini bertujuan untuk menyiapkan data agar lebih terstruktur, konsisten, dan layak digunakan dalam proses data mining. Preprocessing dilakukan melalui beberapa langkah, yaitu pembersihan data, seleksi data, dan penyesuaian format data. Tahapan ini penting karena kualitas data awal sangat memengaruhi hasil klasifikasi yang diperoleh. Data yang sudah melalui preprocessing kemudian digunakan sebagai data latih dalam proses penerapan algoritma.

This work is licensed under a Creative Commons Attribution 4.0

Metode klasifikasi yang digunakan dalam penelitian ini adalah algoritma K-Nearest Neighbor (K-NN). Algoritma ini bekerja dengan cara menghitung jarak antara data uji dan data latih menggunakan metode Euclidean Distance. Setelah jarak dihitung, sistem menentukan sejumlah tetangga terdekat berdasarkan nilai k yang telah ditentukan. Pada penelitian ini, nilai k yang digunakan adalah 15, karena nilai tersebut dipilih berdasarkan hasil evaluasi performa terbaik pada model yang digunakan. Kelas dari data baru kemudian ditentukan berdasarkan mayoritas kelas dari tetangga terdekat.

Pengujian model dilakukan dengan bantuan aplikasi Orange dan perhitungan manual sebagai pembanding. Orange digunakan untuk memproses data mining, membangun model K-NN, serta menampilkan hasil prediksi kualitas buah jeruk bali. Sementara itu, perhitungan manual digunakan untuk menjelaskan tahapan kerja algoritma, mulai dari penentuan nilai k, perhitungan jarak, pengurutan jarak terdekat, hingga penentuan kelas hasil prediksi. Hasil akhir dari proses ini digunakan untuk melihat apakah data prediksi termasuk dalam kategori buah jeruk bali berkualitas baik atau tidak baik.

HASIL DAN PEMBAHASAN

1. Analisis Data dan Karakteristik Variabel Penelitian

Analisis data merupakan tahap awal dalam proses data mining yang bertujuan untuk memahami struktur, karakteristik, dan pola awal dari data yang digunakan. Pada penelitian ini, data yang dianalisis adalah data kualitas buah jeruk bali yang diperoleh berdasarkan sejumlah atribut fisik buah. Data tersebut digunakan sebagai dasar dalam proses klasifikasi menggunakan algoritma K-Nearest Neighbor (K-NN). Melalui tahap analisis data, peneliti dapat mengetahui jumlah data, jenis variabel yang digunakan, serta bentuk label kualitas yang akan diklasifikasikan. Dalam penelitian ini, data yang digunakan terdiri atas 50 data jeruk bali sebagai data latih dengan lima variabel utama, yaitu pigmen kulit buah, kemulusan, kelunakan buah, bobot buah, dan tebal kulit buah. Adapun variabel hasil digunakan sebagai label kelas untuk menunjukkan kualitas buah, yaitu kategori baik dan tidak baik.

This work is licensed under a Creative Commons Attribution 4.0

Tabel 1 Data Jeruk Bali

Data Jeruk Minggu Ke	Pigmen Kulit Buah	Kemulusan	Kelunakan Buah	Bobot Buah	Tebal kulit buah	Hasil
1	0.08	45.9	17.61	1270.88	1.43	1
2	0.05	58.44	18.45	817.25	1.37	2
3	0.03	50.2	22.5	945.25	1.29	2
4	0.08	45.23	24.41	1145.6	1.25	1
5	0.09	47.25	18.25	950.75	1.8	2
6	0.04	35.8	25.4	1300.5	1.45	1
7	0.03	39.92	23.77	850.45	1.2	1
8	0.07	40.25	24.83	780.56	1.27	2
9	0.06	55	18.17	923.49	1.35	1
10	0.08	59.25	17.85	1256.4	1.39	2
11	0.07	38.25	23.81	785.56	1.26	2
12	0.09	56.45	18.82	1236.3	1.26	1
13	0.08	46.93	16.83	1271.85	1.47	1
...	...	...	...	...	...	...
...	...	...	...	...	...	...
47	0.06	43.25	24.83	780.56	1.27	2
48	0.03	39.92	23.77	850.45	1.2	2
49	0.08	43.28	24.83	780.56	1.27	1
50	0.07	44.25	23.83	780.56	1.27	2

Berdasarkan Tabel 1, setiap data jeruk bali memiliki nilai numerik yang mewakili kondisi fisik buah. Variabel pigmen kulit buah menunjukkan kandungan klorofil pada kulit buah, sedangkan kemulusan menggambarkan tingkat kehalusan permukaan buah. Variabel kelunakan buah digunakan untuk melihat tingkat tekstur atau kelembutan buah, bobot buah menunjukkan berat buah, dan tebal kulit buah menunjukkan ukuran ketebalan kulit jeruk bali. Kelima variabel tersebut dipilih karena berhubungan langsung dengan karakteristik fisik yang dapat memengaruhi penilaian kualitas buah.

Selanjutnya, setiap variabel diberikan kriteria tertentu agar data dapat dikelompokkan secara lebih sistematis. Variabel pigmen kulit buah diklasifikasikan berdasarkan nilai klorofil, yaitu nilai kurang dari 0,05 dan lebih dari 0,05. Variabel kemulusan diklasifikasikan berdasarkan batas nilai 40,00%, sedangkan variabel kelunakan buah menggunakan batas nilai 20,00 mm. Untuk variabel bobot buah, batas klasifikasi yang digunakan adalah 1000,00 g, sementara variabel tebal kulit buah diklasifikasikan berdasarkan batas nilai 1,30 cm. Kriteria tersebut digunakan untuk membantu proses penilaian awal terhadap karakteristik buah sebelum data diproses lebih lanjut menggunakan algoritma K-NN.

This work is licensed under a Creative Commons Attribution 4.0

Tabel 2 Variabel Pigmen Kulit Buah

No	Klorofil (mg g -1)	Kriteria
1	<0.05	a
2	>0.05	b

Tabel 3 Variabel Kemulusan

No	Kemulusan(%)	Kriteria
1	<40.00	a
2	>40.00	b

Tabel 4 Variabel Kelunakan Buah

No	Kelunakan (mm 50 g-1 5 detik-1)	Kriteria
1	<20.00	A
2	>20.00	B

Tabel 5 Variabel Bobot Buah (g)

No	Bobot (mg g -1)	Kriteria
1	<1000.00	A
2	>1000.00	B

Tabel 6 Variabel Tebal Kulit (cm)

No	Tebal Kulit(cm)	Kriteria
1	<1.30	A
2	>1.30	B

Selain variabel input, penelitian ini juga menggunakan variabel hasil sebagai label kelas. Label ini menunjukkan kategori kualitas buah jeruk bali berdasarkan data yang telah dikumpulkan. Kategori “baik” diberi kode angka 1, sedangkan kategori “tidak baik” diberi kode angka 2. Pengubahan label kualitas ke dalam bentuk angka dilakukan agar data dapat diproses oleh sistem dan dianalisis menggunakan algoritma klasifikasi. Dengan demikian, seluruh atribut yang digunakan dalam penelitian ini dapat dibaca secara komputasional dan siap digunakan pada tahap pengolahan data berikutnya.

Tabel 7 Variabel Hasil

No	Hasil	Kriteria
1	Baik	1
2	TidakBaik	2

Berdasarkan keseluruhan variabel tersebut, dapat dipahami bahwa kualitas buah jeruk bali dalam penelitian ini ditentukan melalui kombinasi beberapa atribut fisik, bukan hanya berdasarkan satu indikator. Pigmen kulit buah, kemulusan, kelunakan, bobot, dan tebal kulit buah menjadi faktor yang digunakan untuk membedakan buah berkualitas baik

This work is licensed under a Creative Commons Attribution 4.0

dan tidak baik. Data inilah yang kemudian menjadi dasar dalam proses preprocessing dan penerapan algoritma K-Nearest Neighbor pada tahap selanjutnya.

2. Preprocessing Data dan Pembagian Data Uji

Setelah data karakteristik buah jeruk bali dianalisis, tahap berikutnya adalah melakukan preprocessing data. Tahap ini bertujuan untuk menyiapkan data mentah agar lebih terstruktur, konsisten, dan dapat diproses menggunakan algoritma K-Nearest Neighbor. Dalam proses data mining, kualitas data awal sangat memengaruhi hasil klasifikasi yang diperoleh. Oleh karena itu, data perlu diperiksa terlebih dahulu untuk memastikan tidak terdapat kesalahan penulisan, nilai kosong, data ganda, atau format data yang tidak sesuai.

Pada penelitian ini, proses preprocessing dilakukan melalui beberapa tahapan, yaitu data cleaning, data selection, dan penyesuaian format data. Data cleaning dilakukan untuk memastikan data yang digunakan tidak memiliki kesalahan yang dapat mengganggu proses klasifikasi. Data selection dilakukan dengan memilih atribut yang relevan dengan penentuan kualitas buah jeruk bali, yaitu pigmen kulit buah, kemulusan, kelunakan buah, bobot buah, tebal kulit buah, dan hasil kualitas. Selanjutnya, penyesuaian format data dilakukan agar seluruh atribut dapat dibaca dan diproses oleh sistem secara numerik.

Tabel 8 Data Jeruk Bali Preprocessing

Minggu Ke	Pigmen Kulit Buah	Kemulusan	Kelunakan Buah	Bobot Buah	Tebal Kulit buah	Hasil
1	0.08	45.9	17.61	1270.88	1.43	1
2	0.05	58.44	18.45	817.25	1.37	2
3	0.03	50.2	22.5	945.25	1.29	2
4	0.08	45.23	24.41	1145.6	1.25	1
5	0.09	47.25	18.25	950.75	1.8	2
6	0.04	35.8	25.4	1300.5	1.45	1
7	0.03	39.92	23.77	850.45	1.2	1
8	0.07	40.25	24.83	780.56	1.27	2
.....	.....	.....	.....	.....	.....	.....
.....	.....	.....	.....	.....	.....	.....
48	0.03	39.92	23.77	850.45	1.2	2
49	0.08	43.28	24.83	780.56	1.27	1
50	0.07	44.25	23.83	780.56	1.27	2

Berdasarkan Tabel 8, data yang telah melalui proses preprocessing disusun kembali dalam bentuk yang lebih siap digunakan sebagai data latih. Data tersebut tetap memuat atribut utama yang digunakan dalam penelitian, yaitu minggu ke, pigmen kulit buah,

This work is licensed under a Creative Commons Attribution 4.0

kemulusan, kelunakan buah, bobot buah, tebal kulit buah, dan hasil. Pada tahap ini, label kualitas buah sudah dikodekan ke dalam bentuk angka, yaitu 1 untuk kategori "baik" dan 2 untuk kategori "tidak baik". Pengkodean ini diperlukan karena algoritma K-NN melakukan proses klasifikasi berdasarkan perhitungan numerik terhadap jarak antar data.

Setelah data latih disiapkan, tahap berikutnya adalah pembagian data untuk proses pengujian atau prediksi. Pembagian data dilakukan untuk melihat kemampuan model K-NN dalam menentukan kelas kualitas buah pada data baru yang belum memiliki label hasil. Data latih digunakan sebagai acuan dalam proses perhitungan jarak, sedangkan data prediksi digunakan sebagai data uji yang akan diklasifikasikan berdasarkan kedekatannya dengan data latih. Dengan pembagian ini, proses klasifikasi dapat dilakukan secara lebih terarah karena model memiliki data pembanding dan data yang akan diprediksi.

Tabel 9 Data Prediksi

Minggu Ke	Pigmen Kulit Buah	Kemulusan	Kelunakan Buah	Bobot Buah	Tebal Kulit buah	Hasil
1	0.07	32.76	19.61	870.54	1.52	?
2	0.08	49.25	24.74	1274.27	1.15	?

Tabel 9 menunjukkan data prediksi yang terdiri atas dua record buah jeruk bali. Kedua data tersebut memiliki nilai pada masing-masing atribut, yaitu pigmen kulit buah, kemulusan, kelunakan buah, bobot buah, dan tebal kulit buah, tetapi belum memiliki label hasil. Oleh karena itu, kolom hasil pada data prediksi masih diberi tanda "?". Tanda tersebut menunjukkan bahwa kualitas buah belum diketahui dan akan ditentukan melalui proses klasifikasi menggunakan algoritma K-Nearest Neighbor. Data prediksi inilah yang kemudian digunakan pada tahap penerapan algoritma untuk mengetahui apakah buah jeruk bali termasuk kategori baik atau tidak baik.

3. Penerapan Algoritma K-Nearest Neighbor dan Hasil Prediksi

Setelah data melalui tahap preprocessing dan data prediksi telah disiapkan, tahap berikutnya adalah menerapkan algoritma K-Nearest Neighbor (K-NN). Pada penelitian ini, algoritma K-NN digunakan untuk menentukan kualitas buah jeruk bali berdasarkan kedekatan antara data uji dengan data latih. Data latih yang digunakan berjumlah 50 data, sedangkan data uji atau data prediksi berjumlah 2 record. Proses klasifikasi dilakukan

This work is licensed under a Creative Commons Attribution 4.0

dengan membandingkan nilai atribut pada data prediksi terhadap seluruh data latih, yaitu pigmen kulit buah, kemulusan, kelunakan buah, bobot buah, dan tebal kulit buah.

Langkah pertama dalam penerapan algoritma K-NN adalah menentukan nilai k, yaitu jumlah tetangga terdekat yang digunakan sebagai dasar penentuan kelas. Dalam penelitian ini, nilai k yang digunakan adalah 15. Nilai tersebut dipilih karena menghasilkan performa terbaik berdasarkan evaluasi model yang telah dilakukan. Setelah nilai k ditentukan, tahap berikutnya adalah menghitung jarak antara data uji dan data latih. Perhitungan jarak dilakukan menggunakan metode Euclidean Distance, yaitu metode pengukuran jarak yang umum digunakan dalam algoritma K-NN.

$$d(x,y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$$

Rumus tersebut digunakan untuk menghitung kedekatan antara satu data uji dengan masing-masing data latih. Semakin kecil nilai jarak yang dihasilkan, maka semakin dekat hubungan antara data uji dan data latih tersebut. Dalam konteks penelitian ini, data dengan jarak paling kecil dianggap memiliki karakteristik yang paling mirip dengan data jeruk bali yang sedang diuji. Oleh karena itu, hasil perhitungan jarak menjadi dasar utama dalam menentukan kualitas buah jeruk bali.

$$ED1 = \sqrt{\begin{matrix}(0.08-0.07)^2+(45.9-32.76)^2+(17.61-19.61)^2 \\ +(1270.88-870.54)^2+(1.43-1.52)^2\end{matrix}} \quad (1)$$

$$ED1 = \sqrt{(0)+(172.66)+(4)+(160272)+(0.0081)} \quad (2)$$

$$ED1 = \sqrt{160449}$$

$$ED1 = 400.561$$

……hingga perhitungan kebaris 50

Perhitungan jarak dilakukan terhadap seluruh data latih hingga baris ke-50. Setelah semua jarak diperoleh, nilai jarak tersebut kemudian diurutkan dari yang terkecil hingga terbesar. Data dengan jarak terkecil dianggap sebagai tetangga terdekat dari data uji.

This work is licensed under a Creative Commons Attribution 4.0

Karena nilai k yang digunakan adalah 15, maka sistem mengambil 15 data dengan jarak terdekat sebagai dasar penentuan kelas. Selanjutnya, kelas dari data uji ditentukan berdasarkan jumlah kelas terbanyak dari 15 tetangga terdekat tersebut.

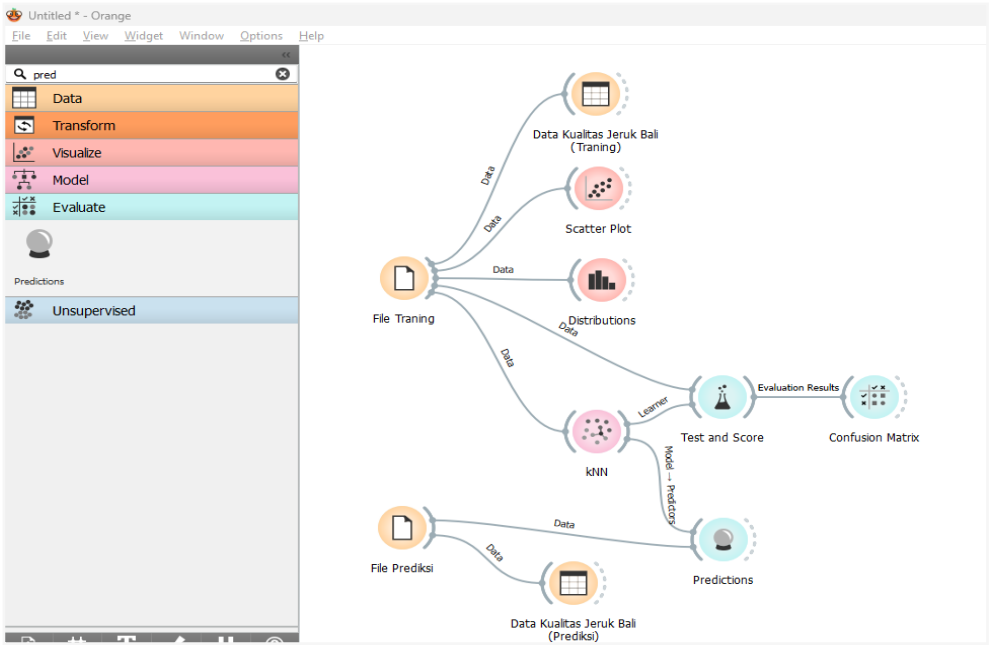
Tabel 10 Hasil Prediksi Perhitungan Manual

Distanc	Rank	Hasil
400.56	11	Baik
59.17	38	Tidak
76.77	36	Tidak
275.38	20	Tidak
81.52	33	Tidak
430.01	6	Baik
21.73	45	Tidak
90.44	26	Tidak
....	....	....
90.81	21	Tidak

Berdasarkan hasil perhitungan manual pada Tabel 10, diperoleh informasi jarak, peringkat atau rank, serta kelas hasil dari setiap data latih terhadap data uji. Dari hasil tersebut, data baru yang diuji menunjukkan bahwa kelas hasil dengan kode 1 atau kategori baik muncul sebanyak 17 data, sedangkan kelas hasil dengan kode 2 atau kategori tidak baik muncul sebanyak 33 data. Namun, dalam proses K-NN, penentuan hasil klasifikasi tidak menggunakan seluruh data, melainkan berdasarkan 15 tetangga terdekat sesuai nilai k yang telah ditentukan. Dari 15 tetangga terdekat tersebut, kelas dengan jumlah terbanyak menjadi dasar penentuan hasil prediksi. Oleh karena itu, data uji diklasifikasikan ke dalam kelas yang memiliki bobot atau jumlah kemunculan paling dominan pada tetangga terdekat.

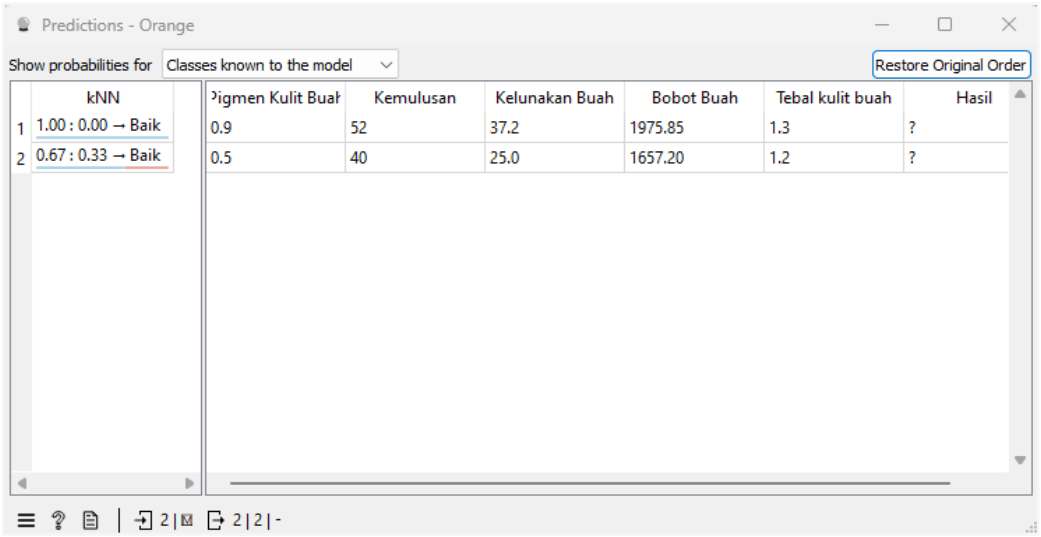
Selain perhitungan manual, penelitian ini juga melakukan pengujian menggunakan aplikasi Orange. Penggunaan Orange bertujuan untuk mempermudah proses data mining, membangun model K-NN, dan menampilkan hasil prediksi secara visual. Pada tahap ini, file data prediksi dihubungkan ke komponen Predictions, kemudian model K-NN yang telah dibuat juga dihubungkan ke komponen yang sama. Proses ini memungkinkan sistem menghitung dan menampilkan hasil klasifikasi berdasarkan model yang digunakan.

This work is licensed under a Creative Commons Attribution 4.0



Gambar 1 Tampilan Prediksi

Gambar 1 menunjukkan proses pengujian data prediksi menggunakan aplikasi Orange. Pada tampilan tersebut, data prediksi dan model K-NN dihubungkan ke fitur Predictions agar sistem dapat melakukan klasifikasi terhadap data baru. Setelah seluruh komponen terhubung, proses prediksi dapat dijalankan untuk mengetahui hasil klasifikasi kualitas buah jeruk bali.



	kNN	Pigmen Kulit Buah	Kemulusan	Kelunakan Buah	Bobot Buah	Tebal kulit buah	Hasil
1	1.00 : 0.00 → Baik	0.9	52	37.2	1975.85	1.3	?
2	0.67 : 0.33 → Baik	0.5	40	25.0	1657.20	1.2	?

Gambar 2 Tampilan Hasil Predictions

This work is licensed under a Creative Commons Attribution 4.0

Berdasarkan Gambar 2, hasil prediksi menunjukkan bahwa dua record data uji yang dimasukkan ke dalam sistem memperoleh hasil klasifikasi "Baik". Artinya, berdasarkan nilai atribut pigmen kulit buah, kemulusan, kelunakan buah, bobot buah, dan tebal kulit buah, kedua data prediksi tersebut memiliki kedekatan karakteristik dengan data latih yang termasuk dalam kategori baik. Hasil ini menunjukkan bahwa algoritma K-NN dapat digunakan untuk membantu proses penentuan kualitas buah jeruk bali secara lebih sistematis berdasarkan data yang tersedia.

Dengan demikian, penerapan algoritma K-Nearest Neighbor dalam penelitian ini dilakukan melalui beberapa tahapan utama, yaitu menentukan nilai k, menghitung jarak menggunakan Euclidean Distance, mengurutkan jarak terdekat, menentukan kelas berdasarkan mayoritas tetangga terdekat, dan melakukan pengujian menggunakan aplikasi Orange. Hasil prediksi menunjukkan bahwa data uji yang digunakan termasuk ke dalam kategori buah jeruk bali berkualitas baik. Proses ini memperlihatkan bahwa K-NN dapat menjadi metode klasifikasi yang sederhana, terukur, dan relevan untuk digunakan dalam penentuan kualitas buah berdasarkan atribut fisik.

SIMPULAN

Berdasarkan hasil analisis data, penelitian ini menggunakan lima variabel utama dalam menentukan kualitas buah jeruk bali, yaitu pigmen kulit buah, kemulusan, kelunakan buah, bobot buah, dan tebal kulit buah. Kelima variabel tersebut digunakan sebagai atribut penentu kualitas, sedangkan variabel hasil digunakan sebagai label kelas dengan kategori "baik" dan "tidak baik". Data yang digunakan berjumlah 50 data latih, kemudian dikodekan dalam bentuk numerik agar dapat diproses oleh sistem klasifikasi. Dengan demikian, karakteristik fisik buah jeruk bali dapat dijadikan dasar dalam proses penilaian kualitas secara lebih terukur.

Pada tahap preprocessing, data disiapkan melalui proses pembersihan, seleksi atribut, dan penyesuaian format data agar layak digunakan dalam proses data mining. Data yang telah melalui tahap ini kemudian digunakan sebagai data latih, sedangkan dua data prediksi digunakan sebagai data uji yang belum memiliki label hasil. Tahapan ini penting karena kualitas data sangat memengaruhi hasil klasifikasi. Data yang sudah tersusun secara konsisten memungkinkan algoritma K-Nearest Neighbor bekerja lebih efektif dalam membandingkan kedekatan antara data uji dan data latih.

This work is licensed under a Creative Commons Attribution 4.0

Penerapan algoritma K-Nearest Neighbor dilakukan dengan menentukan nilai k = 15 dan menghitung jarak antara data uji dengan data latih menggunakan rumus Euclidean Distance. Hasil perhitungan manual dan pengujian menggunakan aplikasi Orange menunjukkan bahwa dua data prediksi yang diuji termasuk dalam kategori kualitas "baik". Hasil evaluasi model juga menunjukkan nilai akurasi sebesar 0,834, classification accuracy sebesar 0,860, F1 sebesar 0,860, precision sebesar 0,861, recall sebesar 0,860, dan MCC sebesar 0,720. Berdasarkan hasil tersebut, algoritma K-NN dapat digunakan sebagai metode klasifikasi untuk membantu menentukan kualitas buah jeruk bali secara lebih sistematis dan berbasis data.

DAFTAR PUSTAKA

[1] N. Fauziah, "Studi epidemiologi penyakit busuk pangkal batang (Phytophthora sp.) dan tingkat ketahanan klon jeruk pamelo (Citrus maxima) lokal di Kabupaten Pangkep," 2022.
[2] Marhawati, "Analisis pengembangan komoditas jeruk pamelo (Citrus grandis L. Osbeck) melalui pendekatan agribisnis di Kabupaten Pangkep," 2019.
[3] A. Pijpe, K. L. M. Gardien, R. E. van Meijeren-Hoogendoorn, E. Middelkoop, and P. P. M. van Zuijlen, "Uji fitokimia dan efektivitas antiinflamasi sediaan gel ekstrak etanol kulit jeruk bali (Citrus maxima) terhadap Propionibacterium acnes pada mencit," in Textbook on Scar Management: State of the Art Management and Emerging Technologies, pp. 109–115, 2025, doi: 10.1007/978-3-030-44766-3_12.
[4] S. A. S. Makrufa, "Uji kualitas fisik dan biologi sediaan gel berbahan dasar ekstrak etanol kulit jeruk bali (Citrus maxima) terhadap hewan coba mencit," 2025.
[5] J. Bagaskara, Teknik Budi Daya Buah Jeruk. Diva Press, 2021.
[6] I. N. Rai, I. N. Wijaya, I. G. N. A. S. Wirya, N. M. Trigunasih, G. Wijana, I. W. Wiraatmaja, I. K. W. Yudha, and I. W. E. K. Utama, Jeruk Keprok Tejakula: Potensi dan Peluang Pemulihannya di Tengah Endemi CVPD dengan Teknologi Budidaya Sehat. Deepublish, 2026.
[7] D. Jollyta, W. Ramdhan, and M. Zarlis, Konsep Data Mining dan Penerapan. Deepublish, 2020.
[8] A. N. Khormarudin, "Teknik data mining: Algoritma K-Means clustering," Jurnal Ilmu Komputer, pp. 1–12, 2016.
[9] S. N. Bakri and L. S. Harahap, "Analisis klasifikasi algoritma K-Nearest Neighbor (K-NN) pada struktur daerah di Kota Medan," Jurnal Ilmu Komputer dan Sistem Informasi, vol. 4, no. 2, pp. 182–193, 2025.
[10] S. T. Fahmi Ruziq, M. R. Wayahdi, and N. Nurhadi, Implementasi Algoritma K-Nearest Neighbor (K-NN) untuk Klasifikasi Data Kesehatan. Hadla Media Informasi, 2026.
[11] T. S. Feri, "Analisis akurasi algoritma K-Nearest Neighbor berbasis Particle Swarm Optimization pada studi kasus diagnosa penyakit jantung," 2025.
[12] N. T. Ujianto, H. Fadillah, A. P. Fanti, A. D. Saputra, and I. G. Ramadhan, "Penerapan algoritma K-Nearest Neighbors (KNN) untuk klasifikasi citra medis," IT-Explore: Jurnal Penerapan Teknologi Informasi dan Komunikasi, vol. 4, no. 1, pp. 33–43, 2025.
[13] A. K. Wardhani, R. I. Sudra, E. Nugraha, and A. N. Putri, "Optimasi nilai K pada algoritma K-Nearest Neighbor untuk klasifikasi kesehatan janin," Journal of Computer Science and Technology (JCS-TECH), vol. 5, no. 1, pp. 44–51, 2025.
[14] S. I. Guslianto, "Klasifikasi kematangan buah sawit berdasarkan fitur warna, bentuk dan tekstur menggunakan algoritma K-NN," JEPIN (Jurnal Edukasi dan Penelitian Informatika), vol. 9, no. 3, pp. 407–414, 2023.
[15] J. G. M. Sanu, "Identifikasi kematangan buah nanas menggunakan citra digital dengan metode K-Nearest Neighbor (K-NN) dan Support Vector Machine (SVM)," HOAQ (High Education of Organization Archive Quality): Jurnal Teknologi Informasi, vol. 16, no. 1, pp. 35–48, 2025.
[16] R. E. Pawening, W. J. Shudiq, and W. Wahyuni, "Klasifikasi kualitas jeruk lokal berdasarkan tekstur dan bentuk menggunakan metode K-Nearest Neighbor (K-NN)," COREAI: Jurnal Kecerdasan Buatan, Komputasi dan Teknologi Informasi, vol. 1, no. 1, pp. 10–17, 2020.

This work is licensed under a Creative Commons Attribution 4.0

[17] K. P. Siwilopo and H. Marcos, "Membandingkan klasifikasi pada buah jeruk menggunakan metode convolutional neural network dan K-nearest neighbor," Komputa: Jurnal Ilmiah Komputer dan Informatika, vol. 12, no. 1, pp. 57–64, 2023.
[18] A. Suryatama Jaya Putra, "Identifikasi kematangan buah jeruk Medan menggunakan metode KNN (K-Nearest Neighbor) berbasis Red, Green, Blue (RGB)," 2023.

This work is licensed under a Creative Commons Attribution 4.0