JDSP
Jurnal Data Science Penusa
E-ISSN: 1234-5678

Implementasi Data Mining Dalam Klasifikasi Sarang Walet yang Berkualitas Berdasarkan Fitur dan Karakteristik Menggunakan Metode K-Nearest Neighbor

Sadarhati Zebua[1)*], Erwin Panggabean[2)*]
Teknologi Informasi, STMIK Pelita Nusantara, Medan, Indonesa[1],
ajasadar352@gmail.com[1]
Received: 30 September 2025 Accepted: 25 Juni 2026 Published: 13 July 2026
Pp. 335-344

***Sadarhati Zebua**

Kata Kunci: Data Mining, K-Nearest Neighbor, Klasifikasi, Sarang Walet, Kualitas

JDSP: Jurnal Data Science Penusa is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).

Abstrak : Sarang burung walet merupakan komoditas bernilai tinggi yang kualitasnya dipengaruhi oleh berbagai fitur dan karakteristik, seperti warna, bentuk, tekstur, serta kandungan mineral. Penilaian kualitas sarang walet secara manual masih bersifat subjektif, membutuhkan waktu lama, serta tidak efisien dalam skala industri. Untuk mengatasi permasalahan tersebut, penelitian ini mengimplementasikan teknik *data mining* dengan metode *K-Nearest Neighbor* (K-NN) dalam proses klasifikasi kualitas sarang walet. Metode K-NN dipilih karena kesederhanaannya serta kemampuannya dalam melakukan klasifikasi berbasis kedekatan jarak antar data. Tahapan penelitian meliputi pengumpulan data, preprocessing, ekstraksi fitur, pembentukan model K-NN, serta evaluasi menggunakan metrik akurasi, presisi, dan recall. Hasil penelitian menunjukkan bahwa metode K-NN mampu mengklasifikasikan sarang walet ke dalam kategori kualitas secara akurat dan konsisten berdasarkan fitur dan karakteristik yang dianalisis. Dengan demikian, penerapan K-NN dalam klasifikasi sarang walet dapat menjadi solusi efektif untuk membantu pelaku industri dalam meningkatkan efisiensi penilaian, mengurangi subjektivitas, serta memperkuat daya saing produk sarang walet di pasar global.

PENDAHULUAN

Sarang burung walet (*Collocalia sp.*) merupakan salah satu komoditas bernilai ekonomi tinggi di Indonesia karena dipercaya memiliki manfaat kesehatan, kecantikan, hingga vitalitas tubuh. Indonesia bahkan menyumbang lebih dari 75% produksi sarang burung walet dunia(Ahmad & -, 2025). Nilai jual sarang walet sangat ditentukan oleh kualitasnya, yang biasanya dilihat dari warna, ukuran, kebersihan, bentuk, serta kerapatan sarang. Namun, proses penilaian kualitas ini masih banyak dilakukan secara manual dan subjektif oleh penjual maupun pembeli, sehingga sering terjadi perbedaan persepsi yang dapat memengaruhi harga jual (Mulyana et al., 2023).

Perbedaan kualitas sarang walet juga dipengaruhi oleh faktor musiman. Pada musim hujan, sarang yang dihasilkan cenderung lebih besar, putih, dan bersih, sedangkan pada musim kemarau hasilnya sebaliknya (Ismail & Yulvianda, 2023). Kondisi ini menimbulkan kebutuhan akan suatu sistem klasifikasi otomatis yang dapat membantu menstandarkan penilaian kualitas sarang walet secara objektif agar pembeli maupun penjual memiliki acuan yang jelas dalam transaksi.

Perkembangan teknologi data mining dan machine learning telah membuka peluang untuk menyelesaikan masalah tersebut. Salah satu metode yang banyak digunakan adalah K-Nearest Neighbor (KNN). Metode KNN dikenal sederhana namun efektif dalam klasifikasi berbasis kedekatan data. Berbagai penelitian telah membuktikan efektivitas KNN, seperti pada klasifikasi mutu telur burung puyuh berdasarkan warna dan tekstur (Ismail & Yulvianda, 2023), identifikasi citra burung Jalak Bali dengan akurasi di atas 98% (Mulyana et al., 2023), serta klasifikasi suara burung menggunakan fitur Mel-Frequency Cepstral Coefficient (Pattraksha et al., 2023). Bahkan, penelitian terbaru menunjukkan bahwa penerapan KNN pada klasifikasi sarang walet dapat mencapai akurasi hingga 90% (Ismail & Yulvianda, 2023).

Selain KNN, metode lain seperti Convolutional Neural Network (CNN) juga pernah digunakan untuk klasifikasi kebersihan sarang walet dengan tingkat akurasi 84%(Ahmad & -, 2025). Akan tetapi, metode tersebut memerlukan komputasi yang lebih kompleks dan data dalam jumlah besar. Oleh karena itu, pemilihan metode KNN pada penelitian ini menjadi relevan karena lebih sederhana, dapat bekerja baik dengan jumlah data terbatas, serta sudah terbukti memberikan hasil klasifikasi yang tinggi pada berbagai domain.

Berdasarkan uraian di atas, terlihat adanya kebutuhan mendesak untuk mengembangkan sistem klasifikasi kualitas sarang walet berbasis data mining yang memanfaatkan metode KNN dengan fitur dan karakteristik citra sarang walet. Penelitian ini diharapkan mampu membantu para peternak dan pelaku usaha dalam menentukan kualitas sarang walet secara cepat, objektif, dan akurat, sehingga dapat meningkatkan nilai jual produk walet di pasar domestik maupun internasional, maka diangkatlah sebuah penelitian untuk penyelesaian masalah dalam

menentukan bahan walet yang baik dengan judul "Implementasi Data Mining Dalam Klasifikasi sarang Walet yang Berkualitas Berdasarkan Fitur dan Karakteristik Menggunakan Metode *K-Nearest Negihbor*".

TINJAUAN LITERATUR

1. Rumusan Masalah

Dari uraian yang telah dijelaskan pada latar belakang, rumusan masalah penelitian dibidang usaha walet sebagai berikut :

1. Bagaimana menerapkan metode KNN dalam menentukan kualitas sarang walet?
2. Bagaimana membangun aplikasi untuk mengklasifikasi sarang walet menggunakan metode K-Nearest Neighbor dengan bahasa pemrograman PHP dan Mysql?

2. Batasan Masalah

Batasan masalah penelitian sebagai berikut :

1. Penelitian ini hanya difokuskan untuk mengklasifikasikan sarang walet berdasarkan kriteria yang ada menggunakan metode KNN.
2. Sistem yang akan dibangun dengan sistem menggunakan bahasa pemrograman PHP dan database MySql.
3. Kriteria yang digunakan dalam penelitian nilai kualitas walet meliputi : warna, ukuran, kebersihan, bentuk, berat/timbangan dan kelembapan.

3. Tujuan Penelitian

Adapun tujuan penelitian bagian dari jawaban rumusan masalah yang telah diteliti ialah sebagai berikut :

1. Menganalisa pengklasifikasian sarang walet agar lebih baik dalam pemilihan sarang walet yang berkualitas baik.
2. Menerapkan data mining dalam klasifikasi sarang walet yang nilainya berkualitas dengan menggunakan K-Nearest Neighbor di bidang usaha walet.
3. Membangun aplikasi sistem teknologi berbasis website untuk memantau dan mengontrol produksi dan pengolahan sarang walet yang berkualitas.

4. Manfaat Penelitian

Manfaat penelitian ialah bagian dari hasil yang diperoleh oleh pneliti di tempat riset yang didapatkan, maka dengan demikian adapun manfaat penelitian yang diperoleh antara lain sebagai berikut :

1. Dengan menerapkan metode K-Nearest Neighbor dapat meningkatkan produksi sarang walet sehingga memperoleh nilai kualitas yang baik
2. Sebagai bahan referensi mahasiswa dalam penelitian untuk membangun sistem teknologi baru yang berbasis *web* dalam pengklasifikasian sarang walet yg berkualitas.

Dapat meningkatkan nilai kinerja usaha walet dan mengembangkan lapangan pekerjaan oleh bidang pembisnis walet terhadap orang orang yg saat ini masih dalam pengangguran.

Data Mining

Data mining bukanlah bidang yang sepenuhnya baru, melainkan hasil perkembangan dari berbagai disiplin ilmu seperti statistik, kecerdasan buatan, dan *machine learning*. Kesulitan mendefinisikannya muncul karena bidang ini mewarisi banyak teknik dari disiplin-disiplin tersebut. Data mining merupakan suatu proses analitis yang bertujuan untuk menggali pengetahuan tersembunyi dari kumpulan data berskala besar. Proses ini dirancang untuk mengatasi keterbatasan pendekatan konvensional dalam menghadapi volume data yang masif, dimensi atribut yang tinggi, serta variasi data yang heterogen. Melalui penerapan teknik statistik, pembelajaran mesin, dan algoritma komputasi, data mining memungkinkan identifikasi pola, tren, dan relasi yang bermakna. Informasi yang dihasilkan kemudian dimanfaatkan untuk mendukung proses pengambilan keputusan berbasis data (data-driven decision making). Proses ini melibatkan tahap-tahap seperti pengumpulan data, pembersihan data, pemilihan atribut, penerapan algoritma, hingga evaluasi hasil. Hasil dari data mining dapat digunakan untuk berbagai keperluan, mulai dari analisis bisnis, prediksi tren, pengelompokan, hingga klasifikasi. Dengan kemampuannya, data mining menjadi salah satu teknologi penting di era big data dan transformasi digital. Dalam penelitian ini, data mining dimanfaatkan untuk mengelompokkan bahan walet berdasarkan fitur dan karakteristiknya secara akurat.

Metode K-Nearest Neighbor

1. Tantangan dalam KNN Classification

Algoritma **K-Nearest Neighbor (KNN)** merupakan salah satu algoritma data mining paling populer karena kesederhanaan dan efektivitasnya. Namun, KNN menghadapi berbagai tantangan, antara lain penentuan nilai *k*, pemilihan tetangga terdekat, pencarian tetangga terdekat, serta aturan klasifikasi (Zhang, 2022). Permasalahan utama adalah kesulitan menentukan nilai *k* yang tepat untuk dataset yang berbeda-beda, karena distribusi data yang tidak seragam. Selain itu, metode pencarian tetangga terdekat membutuhkan komputasi yang besar, terutama pada data berskala besar. Untuk mengatasinya, berbagai metode telah dikembangkan,

seperti adaptasi nilai *k* berdasarkan distribusi data, penggunaan fungsi jarak yang lebih adaptif, dan algoritma pencarian tetangga terdekat secara aproksimasi.

2. Efisiensi KNN untuk Big Data

Menekankan bahwa meskipun KNN sederhana dan banyak diaplikasikan, algoritma ini memiliki kelemahan dalam hal kompleksitas komputasi. Pada dataset besar, pencarian tetangga terdekat memerlukan waktu yang lama karena kompleksitas linear terhadap ukuran data (Zhang et al., 2021). Untuk itu, dikembangkan metode **cluster-based KNN**, di mana data terlebih dahulu dikelompokkan menggunakan algoritma *k-means* atau *landmark-based spectral clustering*, kemudian klasifikasi dilakukan hanya pada cluster terdekat. Hasil penelitian menunjukkan pendekatan ini meningkatkan efisiensi dan akurasi, sehingga KNN dapat diterapkan secara efektif pada big data.

3. Analisis Perbandingan Varian KNN untuk Prediksi data

melakukan studi komparatif terhadap berbagai varian KNN, seperti **Adaptive KNN, Locally Adaptive KNN, Fuzzy KNN, Weighted KNN, Ensemble KNN, Hassanat KNN, dan Generalised Mean Distance KNN**. Hasil penelitian menunjukkan bahwa **Hassanat KNN** memiliki akurasi rata-rata tertinggi (83,62%), diikuti oleh **Ensemble KNN** (82,34%). Studi ini menegaskan bahwa pemilihan varian KNN yang tepat bergantung pada konteks aplikasi dan ukuran dataset (Uddin et al., 2022). Misalnya, varian berbasis fuzzy lebih baik dalam menangani data dengan ketidakpastian, sedangkan varian berbasis ensemble efektif untuk meningkatkan stabilitas hasil klasifikasi.

4. Implementasi KNN pada Analisis Sentimen

Penelitian oleh Isnain dkk. (Isnain et al., 2021) menunjukkan penerapan KNN untuk **analisis sentimen opini publik terkait pembelajaran daring** menggunakan data Twitter. Proses analisis dilakukan melalui *text preprocessing* (case folding, stopword removal, tokenizing, stemming), pembobotan kata menggunakan **TF-IDF**, lalu klasifikasi dengan KNN. Hasil pengujian menunjukkan akurasi terbaik sebesar **84,65% pada k = 10**, dengan presisi 87% dan recall 86%. Hal ini membuktikan bahwa KNN dapat digunakan secara efektif dalam bidang ***Natural Language Processing*** **(NLP),** khususnya analisis sentimen, meskipun nilai *k* yang dipilih sangat memengaruhi hasil akhir.

Prinsip Kerja K-Nearest Neighbors

Algoritma K-Nearest Neighbors (KNN) mengklasifikasikan data baru berdasarkan mayoritas kelas dari K data tetangga terdekat dalam dataset pelatihan. Prosesnya melibatkan penghitungan jarak (biasanya Euclidean) antara data baru dengan data pelatihan, lalu memilih K data terdekat. Kelas yang paling dominan di antara K tetangga tersebut akan menjadi hasil prediksi..

Langkah-langkah dalam K-Nearest Neighbor

Berikut adalah langkah-langkah kerja KNN:

1. Penentuan Jarak: Hitung jarak antara data baru dengan semua data dalam dataset pelatihan. Beberapa metode jarak yang umum digunakan antara lain Euclidean, Manhattan, dan Minkowski.
2. Pemilihan Tetangga Terdekat: Pilih K tetangga terdekat dari data baru berdasarkan jarak terpendek.
3. Penentuan Kelas: Tentukan kelas data baru berdasarkan mayoritas kelas dari K tetangga terdekat. Jika menggunakan regresi, nilai rata-rata dari K tetangga terdekat digunakan.
4. Evaluasi: Lakukan evaluasi pada data uji untuk mengukur kinerja algoritma KNN.

Dalam penerapannya, nilai K perlu ditentukan. Nilai K yang lebih besar akan menghasilkan model yang lebih halus (*smooth*) dan cenderung kurang sensitif terhadap data yang bising, namun mungkin kehilangan detail penting. Sebaliknya, nilai K yang lebih kecil akan lebih sensitif terhadap data yang bising tetapi mungkin lebih mampu menangkap detail penting dalam data. Pemilihan nilai K yang tepat akan bergantung pada dataset yang digunakan dan tujuan dari analisis.

Secara umum, formula Euclidean distance pada 1-*dimentional space* adalah sebagai berikut:

$$dis\ (x_1, x_2) = \sqrt{\sum_{i=0}^{n} (x_{1i} - x_{2i})^2}$$

Keterangan:

dis=nilai euclidian distance data lama dengan data yang baru
n= banyak data
i= iterasi mulai dari data pertama hingga terakhir

Jika hanya ada satu variabel independen, maka jarak dapat dihitung langsung. Namun, jika terdapat banyak variabel, maka semua perbedaan antar variabel dijumlahkan secara kuadrat dan diakar untuk mendapatkan total jarak—misalnya menggunakan rumus **jarak Euclidean multivariat**.

$$dis = \sqrt{\sum_{i=0}^{n}(x_{1i} - x_{2i})^2 + (y_{1i} - y_{2i})^2 + \cdots}$$

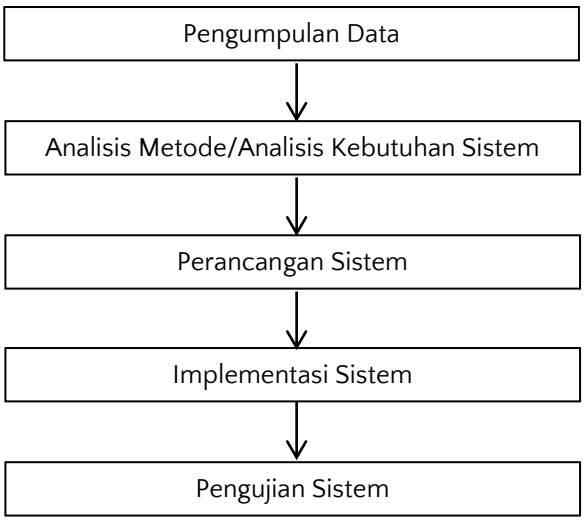
METODOLOGI PENELITIAN

Penelitian ini termotivasi dari permasalahan proses penilaian kualitas yang masih banyak dilakukan secara manual dan subjektif oleh penjual mapun pembeli, sehingga terjadi persepsi yang dapat memengaruhi harga jual.

Perkembangan teknologi data mining dan machine learning telah membuka peluang baru untuk mengatasi berbagai permasalahan, salah satunya melalui penggunaan metode **K-Nearest Neighbor (KNN)**. Metode KNN merupakan algoritma pembelajaran mesin yang melakukan klasifikasi atau prediksi data baru berdasarkan kedekatannya dengan data latih yang sudah ada. Algoritma ini bekerja dengan mencari **K** data terdekat terhadap data uji, kemudian menentukan kelas data baru berdasarkan mayoritas kelas dari tetangga terdekat tersebut.

1. Kerangka Kerja Penelitian

Kerangka kerja merupakan rangkaian tahapan yang dijalankan dalam proses penyelesaian masalah yang akan dibahas. Kerangka penelitian ini berfungsi sebagai panduan yang membantu dalam penyusunan dan pelaksanaan penelitian secara sistematis. Berikut kerangka kerja yang akan digunakan adalah :



Gambar 1. Kerangka Kerja Penelitian

2. Uraian Kerangka Kerja

Berdasarkan Kerangka Kerja Penelitian Pada Gambar 1 Kerangka Kerja Penelitian Dapat Diuraiakan Sebagai Berikut:

a. Pengumpulan Data

Pengumpulan data dilakukan dengan mengamati permasalahan dalam menentukan sarang walet yang berkualitas berdasarkan kriteria sebagai berikut :

1. Warna.
2. Berat.
3. Kerapatan serat dan susut.
4. Kadar air/Kelembaban.
5. Kebersihan
6. Bentuk
7. Ukuran

b. Analisis Metode/Analisis Kebutuhan Sistem

Tahapan ini mencakup analisis terhadap masalah yang diperoleh selama penelitian berdasarkan hasil observasi dan wawancara, sehingga dapat ditemukan solusi yang tepat. Penelitian ini membahas dan menyajikan uraian hasil analisis dalam menentukan sarang walet berkualitas dengan menggunakan metode *K-Nearest Neighbor*.

c. Perancangan Sistem

Pada tahap ini, penelitian difokuskan pada perancangan sistem dengan menggunakan pemodelan UML (Unified Modeling Language) yang bertujuan untuk menggambarkan alur penyelesaian masalah secara jelas.

Diagram UML yang digunakan dalam tahap ini meliputi Use Case Diagram, Activity Diagram, dan Class Diagram.

Model yang dirancang dengan beberapa tahap yaitu:

1. Pemilihan fitur utama yang berpengaruh terhadap kualitas sarang walet.
2. Implementasi metode K-Nearest Neighbor dengan kriteria.

d. Implementasi Sistem

Pada tahap ini akan dibahas terkait tentang aplikasi *development tools* yang akan digunakan untuk membangun sistem. Aplikasi sistem yang digunakan dalam penelitian ini yaitu dengan menggunakan aplikasi sistem teknologi berbasis *web*.

Dengan membangun sebuah aplikasi sistem ini, maka dapat membantu dan mengontrol produksi dan pengolahan sarang walet yang berkualitas. Maka dari itu, dengan adanya sistem aplikasi ini bisa diterapkan di dalam Perusahan di tempat penelitian sebagai contoh masukan untuk membantu dalam menjalankan usaha walet dengan efisien.

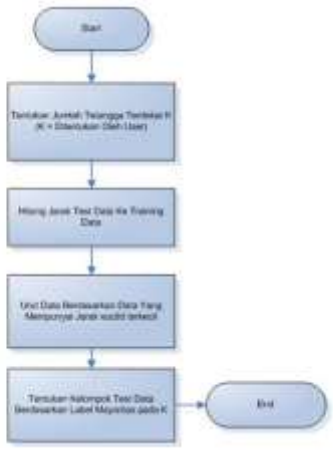
e. Pengujian Sistem

Pada tahapan ini, secara singkat terkait teknik yang digunakan dalam pengujian sistem yaitu dengan menggunakan *Blackbox Testing*. Pengujian ini dilakukan untuk memperbaiki kinerja dari sistem dengan cepat.

Analisis Sistem yang dibangun

Kegiatan analisis sistem yang dibangun adalah untuk menentukan kelompok sarang walet yang baru saja masuk ke dalam sistem. Tujuan dari tahap analisis ini adalah memahami dengan baik terhadap sistem yang sedang berjalan dan sistem yang akan dibangun, yaitu :

Berdasarkan masalah diatas, solusi untuk menganalisis yaitu dengan merancang sebuah sistem yang mampu membuat klasifikasi terhadap sarang walet yang baru saja dimasukkan ke dalam sistem. Tahapan sistem yang akan dibangun dapat dilihat dari gambar flowchat seperti berikut :



Gambar 2. Flowchart Sistem yang akan dibangun

Gambar flowchart diatas menjelaskan:

a. admin menentukan jumlah tetangga terdekat yang disebut dengan K. K dalam penelitian ini menggunakan K=7. Artinya data test akan diuji ke dalam 7 data training yang lebih dekat dengannya.
b. Menghitung jarak ecludian data test dengan data training yang ada
c. Mengurutkan data training berdasarkan nilai ecludian dengan nilai ecludian data test.
d. Menentukan kelompok data test dengan melihat jumlah kelompok data training terdekat yang mayoritas berdasarkan nilai ecludiannya.

Analisis Algoritma *K-Nearest Neighbor*

Salah satu perhitungan yang dapat digunakan untuk klasifikasi data dalam data mining adalah algoritma *K-Nearest Neighbor*. Metode ini digunakan untuk membandingkan data yang baru masuk dengan menganalisa data yang sudah ada di dalam database (data training) dengan mencari nilai euclidean ke setiap data, kemudian ditentukan nilai K atau jumlah tetangga atau data terdekat dengan data baru. Misalkan K=7 maka, kita mencari 7 data euclidean terkecil, kemudian menghitung jumlah data yang terdapat pada kelompok mayoritas. Jika sudah ditemukan, maka data baru tersebut dimasukkan ke dalam kelompok data yang mayoritas.

Berikut adalah data yang akan digunakan untuk penerapan algoritma *K-Nearest Neighbor* dalam sistem yang akan dibangun.

Tabel 1
Data training sarang walet

No	kebersihan	ukuran_cm	berat_gr	kualitas

1	60	2	5	Kualitas Buruk
2	50	1	3	Kualitas Buruk
3	70	3	10	Kualitas Buruk
4	80	4	15	Kualitas Baik
5	75	5	12	Kualitas Buruk
6	90	8	10	Kualitas Baik
7	85	6	15	Kualitas Baik
8	90	7	10	Kualitas Baik
9	60	8	5	Kualitas Buruk
10	70	6	3	Kualitas Buruk
11	98	1	3	Kualitas Baik
12	50	2	5	Kualitas Buruk
13	60	4	6	Kualitas Buruk
14	95	6	7	Kualitas Baik
15	85	8	1	Kualitas Baik
16	80	6	3	Kualitas Baik
17	60	3	5	Kualitas Buruk
18	70	5	6	Kualitas Buruk
19	75	1	3	Kualitas Buruk
20	60	2	8	Kualitas Buruk

Tabel di atas adalah data yang akan digunakan sebagai data training untuk menentukan kelompok/klasifikasi data test yang akan dimasukkan.

Penerapan Algoritma K-Nearest Neighbor

Berikut adalah langkah-langkah klasifikasi data dengan nilai variabel yang ada menggunakan metode *K-Nearest Neighbor*.

Langkah 1:

Tentukan nilai K (jumlah data yang terdekat atau dengan nilai euclidean yang terkecil terhadap data baru).

Misal: data baru yang masuk adalah sebagai berikut:

Tabel 2
Sampel data walet baru

Id	Kebersihan	ukuran_cm	berat_gr
1	75	3	4

Tabel di atas ada sampel data test walet baru yang akan digunakan untuk menguji klasifikasi menggunakan KNN dengan data training di tabel 4.1.

Cari nilai euclidian data baru dengan data training

d(1,1)= $\sqrt{(60-75)^2+(2-3)^2+(5-4)^2} = 15.06651$

d(2,1)= $\sqrt{(50-75)^2+(1-3)^2+(3-4)^2} = 25.09980$

dan seterusnya sampai hingga data training yang ke 20. Dari perhitungan diatas di dapat nilai euclidean setiap data training sebagai berikut.

Tabel 3
Nilai euclidean setiap data training

No	Kebersihan	Ukuran (cm)	Berat (gr)	Kualitas	Euclidean distance
1	60	2	5	Kualitas Buruk	15.06651917
2	50	1	3	Kualitas Buruk	25.0998008
3	70	3	10	Kualitas Buruk	7.810249676
4	80	4	15	Kualitas Baik	12.12435565
5	75	5	12	Kualitas Buruk	8.246211251
6	90	8	10	Kualitas Baik	16.91153453
7	85	6	15	Kualitas Baik	15.16575089
8	90	7	10	Kualitas Baik	16.64331698
9	60	8	5	Kualitas Buruk	15.84297952
10	70	6	3	Kualitas Buruk	5.916079783
11	98	1	3	Kualitas Baik	23.10844002
12	50	2	5	Kualitas Buruk	25.03996805
13	60	4	6	Kualitas Buruk	15.16575089
14	95	6	7	Kualitas Baik	20.4450483

15	85	8	1	Kualitas Baik	11.5758369
16	80	6	3	Kualitas Baik	5.916079783
17	60	3	5	Kualitas Buruk	15.03329638
18	70	5	6	Kualitas Buruk	5.744562647
19	75	1	3	Kualitas Buruk	2.236067977
20	60	2	8	Kualitas Buruk	15.55634919

Tabel di atas adalah hasil perhitungan ecludian setiap data training terhadap data test yang sudah dimasukkan yang ada pada tabel 2.

HASIL PENELITIAN

Langkah berikutnya adalah mendeteksi data dengan nilai terendah sebanyak 7 data karena ditentukan dari awal jumlah K adalah 7. Maka didapatlah tabel sebagai berikut.

Tabel 4
Data training dengan nilai K=7

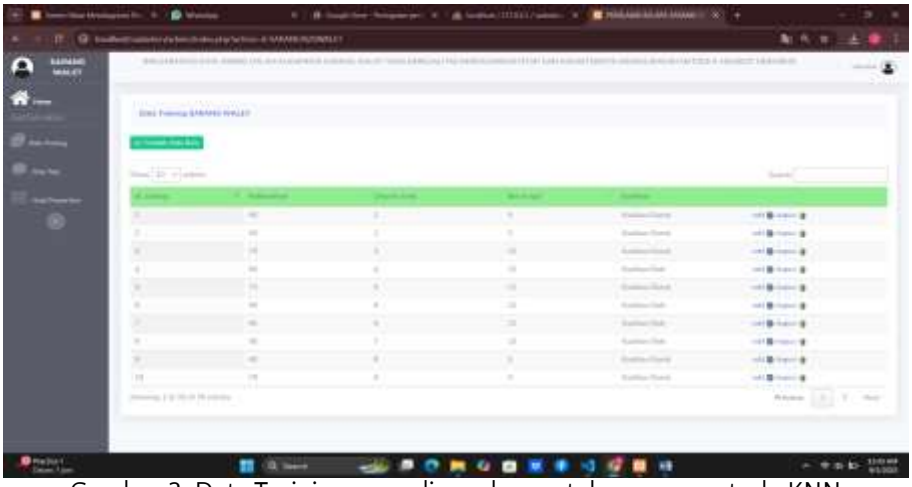
No	kebersihan	Ukuran(cm)	Berat(gr)	Kualitas	Euclidean distance	K=7
1	60	2	5	Kualitas Buruk	15.06651917	10
2	50	1	3	Kualitas Buruk	25.0998008	20
3	70	3	10	Kualitas Buruk	7.810249676	5
4	80	4	15	Kualitas Baik	12.12435565	8
5	75	5	12	Kualitas Buruk	8.246211251	6
6	90	8	10	Kualitas Baik	16.91153453	16
7	85	6	15	Kualitas Baik	15.16575089	11
8	90	7	10	Kualitas Baik	16.64331698	15
9	60	8	5	Kualitas Buruk	15.84297952	14
10	70	6	3	Kualitas Buruk	5.916079783	3
11	98	1	3	Kualitas Baik	23.10844002	18
12	50	2	5	Kualitas Buruk	25.03996805	19
13	60	4	6	Kualitas Buruk	15.16575089	11
14	95	6	7	Kualitas Baik	20.4450483	17
15	85	8	1	Kualitas Baik	11.5758369	7
16	80	6	3	Kualitas Baik	5.916079783	4
17	60	3	5	Kualitas Buruk	15.03329638	9
18	70	5	6	Kualitas Buruk	5.744562647	2
19	75	1	3	Kualitas Buruk	2.236067977	1
20	60	2	8	Kualitas Buruk	15.55634919	13

Data pada tabel diatas menunjukkan data-data training serta menandai 7 data terdekat berdasarkan nilai ecludian yang sudah dihitung dengan data test yang dimasukkan.

Dari data diatas dapat kita lihat bahwa data tetangga 7 terdekat dengan data baru adalah data ke 3, 5, 10, 15, 16, 18, 19. Dari 7 data yang terdekat ini maka dapat dibandingkan atau dicari data mayoritas, yaitu ada 5 data dengan kualitas “Kualitas Buruk” dan 2 data dengan kualitas “Kualitas Baik”. Dengan demikian dapat disimpulkan bahwa data yang baru (data test) masuk ke dalam kelompok “Kualitas Buruk”.

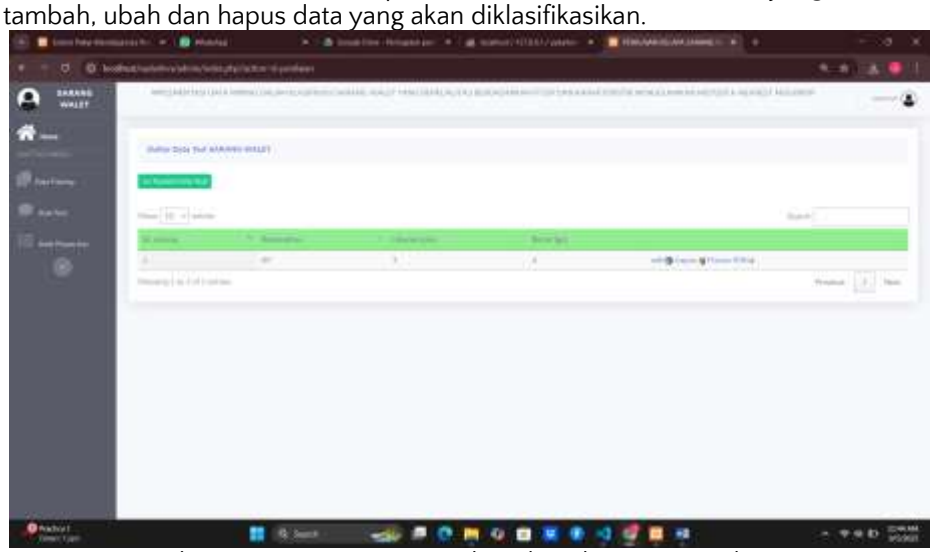
Implementasi adalah prosedur untuk merealisasikan sistem berdasarkan dokumen rancangan yang telah disetujui dan diuji, dengan cara menginstal serta mulai menggunakan sistem baru yang telah disempurnakan. Tahap ini bertujuan untuk memastikan bahwa sistem atau aplikasi yang dirancang benar-benar dapat mencapai tujuan yang telah ditetapkan.

Halaman data training adalah halaman yang menampilkan data training serta tombol operasi seperti tambah data, edit dan hapus data training sarang walet.



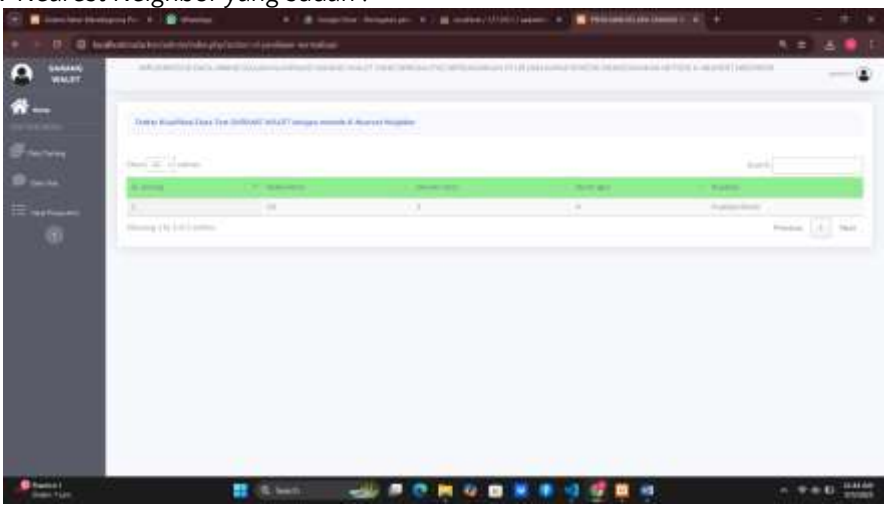
Gambar 3. Data Training yang digunakan untuk proses metode KNN

Halaman data test adalah untuk menampilkan data nilai kriteria alternatif yang sudah dimasukkan serta operasi seperti tambah, ubah dan hapus data yang akan diklasifikasikan.



Gambar 4. Data Testing yang akan diuji dengan metode KNN

Halaman hasil proses K-Nearest Neighbor adalah untuk menampilkan data test yang sudah di proses dengan metode K-Nearest Neighbor yang sudah .



Gambar 5. Hasil pengolahan data testing dengan metode KNN

PEMBAHASAN

Pengujian sistem ini digunakan dengan menggunakan metode Black Box testing untuk mengetahui kinerja sistem yang sudah dibangun.

Tabel 5
Hasil pengujian sistem menggunakan metode Black Box Testing

No	Fitur yang diuji	Skenario Pengujian	Hasil yang diharapkan	Hasil Aktual	Status
1	Form Login	Pengguna memasukkan username dan password yang valid	Sitem menampilkan menu utama sistem	Selesai	Lulus
2	Menu Utama	Pengguna berhasil masuk dan melihat semua menu pada system	Seluruh menu tampil	Selesai	Lulus
3	Menu Data training	Menampilkan data training sarang walet, menambah data training sarang walet baru	Data training sarang walet tersimpan ke dalam database	Selesai	Lulus
4	Menu data test sarang wallet	Menampilkan data test dan memasukkan data test sarang walet yang baru	Data nilai tersimpan ke dalam database	Selesai	Lulus
5	Menu Proses KNN	Menjalankan proses klasifikasi dengan menggunakan metode K-NN	Data sarang walet menghasilkan nilai dan tersimpan ke dalam database	selesai	Lulus

Tabel di atas menjelaskan tetang pengujian sistem yang dilakukan dengan metode black box testing untuk mengetahui apakah seluruh menu pada sistem berjalan dengan baik.

Kelebihan dan Kelemahan Sistem

Adapun kelebihan dan kelemahan aplikasi Implementasi Data Mining dalam Klasifikasi Sarang Walet Yang Berkualitas Berdasarkan Fitur Dan Karakteristik Menggunakan Metode K-Nearest Neighbor ini dapat dilihat pada sub di bawah ini.

Kelebihan sistem

Kelebihan Dari Implementasi Data Mining Dalam Klasifikasi Sarang Walet Yang Berkualitas Berdasarkan Fitur Dan Karakteristik Menggunakan Metode K-Nearest Neighbor yaitu:

1. Sistem ini menampilkan hasil berupa tabel yang berisi data-data sarang walet yang dianggap sesuai dikategorikan sesuai dengan klasifikasi yang benar berdasarkan data training yang ada.
2. Sistem memungkinkan penambahan data uji dan data pelatihan sesuai dengan kebutuhan.
3. Sistem memiliki desain yang user-friendly sehingga memudahkan pengguna dalam penggunaannya.

Kelemahan Sistem

Kelemahan dari Implementasi Data Mining Dalam Klasifikasi Sarang Walet Yang Berkualitas Berdasarkan Fitur dan Karakteristik Menggunakan Metode K-Nearest Neighbor yaitu:

1. Hasil akhir K-Nearest Neighbor adalah secara otomatis dapat menentukan klasifikasi data test dengan mempelajari data training yang ada.
2. Sistem yang dibuat masih bersifat offline atau hanya dapat dijalankan di komputer stand alone karena belum di hosting di internet.

KESIMPULAN

Dari hasil pembuatan tugas akhir implementasi data mining dalam klasifikasi sarang walet yang berkualitas berdasarkan fitur dan karakteristik menggunakan metode k-nearest neighbor, dapat ditarik beberapa kesimpulan antara lain:

1. Berdasarkan hasil pengolahan data dan analisa yang telah dilakukan dapat disimpulkan bahwa sarang walet diklasifikasikan dengan mempelajari data training terhadap data testnya, maka akan dihasilkan klasifikasi yang baik.

2. Sistem yang dibuat menampilkan klasifikasi data sarang walet berdasarkan kriteria dan data training sehingga mempermudah pengklasifikasian sarang walet.

Saran

Saran yang dapat diberikan untuk pengembangan lebih lanjut adalah:

1. Diharapkan klasifikasi untuk data sarang walet di perusahaan menggunakan ketentuan yang sudah dihasilkan oleh sistem yang sudah dibangun ini.
2. Untuk penelitian selanjutnya dapat dikembangkan menggunakan android.

REFERENCES

Ahmad, S., & -, S. (2025). Analisis Cnn Untuk Klasifikasi Kebersihan Sarang Burung Walet. *Jurnal Informatika Dan Teknik Elektro Terapan*, *13*(3). https://doi.org/10.23960/jitet.v13i3.6685

Ismail, M., & Yulvianda, R. (2023). Penerapan Data Mining Dengan Metode K-Nearest Neighbor Terhadap Klasifikasi Sarang Walet. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, *7*(3), 1220. https://doi.org/10.30865/mib.v7i3.6431

Isnain, A. R., Supriyanto, J., & Kharisma, M. P. (2021). Implementation of K-Nearest Neighbor (K-NN) Algorithm For Public Sentiment Analysis of Online Learning. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, *15*(2), 121. https://doi.org/10.22146/ijccs.65176

Mulyana, D. I., Zuhari, A. B., & Yel, M. B. (2023). Klasifikasi Citra Burung Jalak Bali dengan Menggunakan Algoritma K-Nearest Neighbor (KNN). *Jurnal Teknik Elektro Dan Komputasi (ELKOM)*, *5*(1), 57–67. http://jurnal.unmuhjember.ac.id/index.php/ELKOM/article/view/9337

Pattraksha, A. A. G. R. K., Supriana, I. W., Widiartha, I. M., Rahning Putri, L. A. A., Gede Dwidasmara, I. B., & Atmaja Darmawan, I. D. M. B. (2023). Classification of Bird Sounds Using the Mel-Frequency Cepstral Coefficient (MFCC) and K-Nearest Neighbor (KNN) Methods. *JELIKU (Jurnal Elektronik Ilmu Komputer Udayana)*, *11*(3), 679. https://doi.org/10.24843/jlk.2023.v11.i03.p25

Purnamasari, Y., Pudjiantoro, T. H., & Nursantika, D. (2017). Sistem Penilaian Kinerja Dosen Teladan Menggunakan Metode Simple Multy Attribute Rating Technique (Smart). *Jurnal Teknologi Elektro*, *8*(1). https://doi.org/10.22441/jte.v8i1.1372

Ramdany, S. (2024). Penerapan UML Class Diagram dalam Perancangan Sistem Informasi Perpustakaan Berbasis Web. In *Journal of Industrial and Engineering System* (Vol. 5, Issue 1, pp. 7–10). https://doi.org/10.31599/2e9afp31

Rizki Prayogo, P., & Hadi Susilo, P. (2022). Sistem Pendukung Keputusan dalam Menentukan Kualitas Sarang Burung Walet Terbaik Menggunakan Metode Simple Additive Weighting (SAW). *Insearch: Information System Research Journal*, *2*(02), 83–89. https://doi.org/10.15548/isrj.v2i02.4363

Uddin, S., Haque, I., Lu, H., Moni, M. A., & Gide, E. (2022). Comparative performance analysis of K-nearest neighbour (KNN) algorithm and its different variants for disease prediction. *Scientific Reports*, *12*(1), 1–11. https://doi.org/10.1038/s41598-022-10358-x

Zhang, S. (2022). Challenges in KNN Classification. *IEEE Transactions on Knowledge and Data Engineering*, *34*(10), 4663–4675. https://doi.org/10.1109/TKDE.2021.3049250

Zhang, S., Deng, Z., Cheng, D., Zong, M., & Zhu, X. (2021). Author's Accepted Manuscript Efficient kNN Classification Algorithm for Big Data Efficient kNN Classification Algorithm for Big Data. *Neurocomputing*. http://dx.doi.org/10.1016/j.neucom.2015.08.112%5Cnwww.elsevier.com/locate/neucom