

# PREDIKSI KUALITAS GENTENG MANTILI BERDASARKAN KOMPOSISI BAHAN BAKU MENGGUNAKAN ALGORITMA K-NEAREST NEIGHBOR

Yusuf Bahtiar\*<sup>1)</sup>, Joni Maulindar<sup>2)</sup>, Margaretha Evi Yuliana<sup>3)</sup>

1. Universitas Duta Bangsa Surakarta, Indonesia
2. Universitas Duta Bangsa Surakarta, Indonesia
3. Universitas Duta Bangsa Surakarta, Indonesia

## Article Info

**Kata Kunci:** Data Mining; Genteng Mantili; K-Nearest Neighbor; Klasifikasi

**Keywords:** Classification; Data Mining; Mantili Tile; K-Nearest Neighbor

## Article history:

Received 29 September 2024

Revised 13 Oktober 2024

Accepted 4 November 2024

Available online 4 December 2024

## DOI :

<https://doi.org/10.29100/jupi.v9i4.5589>

\* Corresponding author.

Yusuf Bahtiar

E-mail address:

[202030292@mhs.udb.ac.id](mailto:202030292@mhs.udb.ac.id)

## ABSTRAK

Permasalahan yang dialami pada industri genteng saat ini adalah fluktuasi harga bahan baku dan ketersediaan bahan baku yang tidak menentu. Proses penentuan komposisi yang tepat melalui uji coba manual memerlukan waktu dan sumber daya yang dapat menghabiskan tenaga dan biaya. Oleh karena itu, diperlukan sebuah sistem yang dapat memprediksi kualitas genteng mantili. Tujuan penelitian ini untuk memprediksi kualitas genteng mantili berdasarkan bahan baku menggunakan algoritma *K-Nearest Neighbor* (K-NN). Komponen bahan baku yang digunakan yaitu tanah liat lempung, tanah liat hitam, tanah liat merah, tanah liat padas, tanah pasir sungai, dan kaolin. Klasifikasi prediksi yang digunakan dalam penelitian ini adalah "bagus" dan "tidak bagus". Data yang digunakan sebanyak 36, kemudian dipisahkan menjadi 26 data latih dan 10 data uji. Penelitian dimulai dengan memodelkan perhitungan K-NN menggunakan *RapidMiner*, kemudian dilanjutkan dengan mengembangkan sistem menggunakan bahasa pemrograman PHP dengan basis data MySQL. Pengujian pertama dilakukan dengan nilai  $k=5$  menggunakan 26 data latih dan 10 data uji, didapatkan hasil nilai akurasi sebesar 80% dan nilai *error* mencapai 20%. Pengujian kedua dilakukan dengan memasukkan data aktual yang diperoleh di lapangan menunjukkan hasil prediksi yang sesuai. Oleh karena itu, dapat disimpulkan bahwa sistem ini dapat digunakan untuk memprediksi kualitas genteng mantili.

## ABSTRACT

The problems experienced in the roof tile industry at this time are fluctuations in the price of raw materials and the uncertain availability of raw materials. The process of determining the right composition through manual trials requires time and resources that can be labor-intensive and costly. Therefore, a system is needed that can predict the quality of mantili roof tiles. The purpose of this research is to predict the quality of mantili roof tiles based on raw materials using the K-Nearest Neighbor (K-NN) algorithm. The raw material components used are loamy soil, black clay, red clay, silt loam, river sand, and kaolin. The prediction classifications used in this study are "good" and "not good". The data used was 36, then separated into 26 training data and 10 testing data. The research began by modeling the K-NN calculation using *RapidMiner*, then continued by developing the system using the PHP programming language with the MySQL database. The first test was conducted with a value of  $k = 5$  using 26 training data and 10 testing data, obtained an accuracy value of 80% and an error value of 20%. The second test was conducted by inputting actual data obtained at the research site shows the appropriate prediction results. Therefore, it can be concluded that this system can be used to predict the quality of mantili roof tiles.

## I. PENDAHULUAN

**G**ENTENG adalah materi yang digunakan untuk atap rumah dalam proses pembangunan sebuah rumah [1]. Genteng digunakan sebagai pelindung penghuni rumah dari panas dan hujan [2]. Genteng dibuat dari tanah yang telah diproses melalui pembakaran dan dibentuk menjadi bentuk persegi panjang dengan ketebalan dan lekukan tertentu [3]. Meskipun terkesan tradisional, genteng dari tanah liat masih sangat diminati dikarenakan tahan cuaca dan minim perawatan [4].

Penelitian ini dilaksanakan di Dukuh Kebak RT 001 RW 013 Desa Wirun Kecamatan Mojolaban Kabupaten Sukoharjo dengan narasumber seorang pengrajin genteng mantili bernama Bapak Sigit Purnomo. Industri genteng menghadapi banyak masalah seiring berjalannya waktu. Biaya bahan baku dapat mengalami fluktuasi sejalan dengan perubahan kondisi ekonomi atau volume produksi pada suatu periode tertentu [2]. Bahan baku yang tidak tersedia dapat menghambat produksi dan menyebabkan keterlambatan dalam memenuhi pesanan [5]. Proses penentuan komposisi yang tepat melalui uji coba manual memerlukan waktu dan sumber daya yang dapat menghabiskan tenaga dan biaya [6].

Masalah yang telah dipaparkan di atas dapat diselesaikan dengan penerapan sistem prediksi kualitas genteng mantili. Dampak perkembangan teknologi dan komunikasi juga turut memfasilitasi akses sistem ini [7]. Sistem yang dibuat dirancang berbasis web untuk memungkinkan akses dari mana saja dan kapan saja [8]. Sistem ini melakukan prediksi mengenai kualitas genteng berdasarkan komposisi bahan bakunya, seperti tanah liat lempung, tanah liat hitam, tanah liat merah, tanah liat padas, tanah pasir sungai, dan kaolin. Atribut label sebagai target prediksi dalam penelitian ini adalah bagus dan tidak bagus. Beberapa kriteria genteng yang baik dan berkualitas meliputi kekuatan, presisi, ketahanan retak, dan warna merah cerah [9]. Dalam melakukan prediksi hasil kualitas genteng mantili, algoritma yang diterapkan untuk prediksi adalah *K-Nearest Neighbor (K-NN)* [10]. Pembangunan sistem menggunakan algoritma K-NN bertujuan untuk menentukan kelompok data dengan cara menghitung perbandingan jarak terdekat antara data bahan baku yang akan diuji dengan data bahan baku lain yang telah menjadi data latih. Nilai mayoritas dari parameter  $k$  digunakan sebagai penentu akhir klasifikasi data uji terkait kualitas genteng mantili, apakah termasuk dalam kelompok yang bagus atau tidak bagus.

Penelitian yang dilakukan pada tahun 2022 berfokus pada kualitas genteng mantili [11]. Penelitian dilakukan di Desa Wirun dan melibatkan perbandingan genteng sokka, mantili, mini, dan ABL. Menggunakan perhitungan *Multi Attribute Utility Theory (MAUT)*, genteng mantili menunjukkan kualitas unggul dan memiliki nilai tertinggi 9,32. Namun, atribut spesifik dari bahan baku yang digunakan tidak dijelaskan secara menyeluruh. Oleh karena itu, penelitian ini bertujuan untuk mengeksplorasi hubungan antara kualitas genteng mantili dan komposisi bahan baku secara lebih rinci.

Beberapa penelitian terdahulu yang berkaitan dengan klasifikasi menggunakan K-NN antara lain, melakukan klasifikasi objek baru berdasarkan label mayoritas dari kedekatan jarak antara kedua data berdasarkan jumlah nilai  $k$  yang digunakan [12]. Akurasi K-NN menjadi paling tinggi ketika digunakan untuk analisis sentimen mengenai relokasi Ibukota Nusantara, mencapai tingkat akurasi 88,12% dibandingkan dengan akurasi *Naive Bayes* sebesar 82,27% [13]. Dalam mendiagnosa penyakit jantung, K-NN dengan metode jarak *Euclidean* memiliki tingkat akurasi 65%, sedangkan menggunakan *Naive Bayes* hanya 58% [14]. Algoritma K-NN menunjukkan kinerja yang baik dalam mengklasifikasikan tingkat kematangan buah nenas sesuai dengan warnanya [15], meramalkan penjualan obat di apotek yang terletak di Palembang [16], dan efektif dalam mengklasifikasikan data kelayakan pengajuan kartu kredit [17].

Maka berdasarkan pendahuluan di atas, penelitian ini akan melakukan prediksi kualitas genteng mantili berdasarkan komposisi bahan baku dengan algoritma *K-Nearest Neighbor*. Selain itu, penelitian ini akan mengembangkan sistem yang menggunakan hasil dari penerapan algoritma untuk memprediksi kualitas genteng mantili berdasarkan komposisi bahan bakunya. Diharapkan aplikasi ini dapat meningkatkan efisiensi produksi genteng mantili, sehingga pengrajin dapat memperoleh keuntungan yang lebih besar.

## II. METODE PENELITIAN

### A. Teknik Pengumpulan Data dan Pemrosesan Data

Metode yang diterapkan dalam pengumpulan data adalah melalui proses wawancara. Wawancara dilakukan untuk mengetahui bahan baku yang digunakan, yaitu tanah liat lempung, tanah liat hitam, tanah liat merah, tanah liat padas, tanah pasir sungai, dan kaolin.

Data yang diperoleh dari proses wawancara terdiri dari 36 data. Data tersebut masih dalam bentuk yang mendeskripsikan komposisi bahan baku, dalam pemrosesan data awal ini akan dilakukan pemberian kriteria, sub-kriteria, dan nilai pembobotan pada sub-kriteria [18]. Pemrosesan awal data dilakukan perlu dilakukan karena data

hasil wawancara tidak bertipe numerik. Diperlukan transformasi data dari kualitatif ke kuantitatif karena data awal bukan numerik [19].

Dalam penelitian ini, digunakan beberapa kriteria untuk menganalisis komposisi bahan baku genteng mantili. Kriteria tersebut dipilih untuk mempersingkat penulisan variabel. Setiap kriteria memiliki sub-kriteria yang disesuaikan dengan jumlah bahan baku yang digunakan. Untuk tanah liat lempung, hitam, merah, dan padas, penentuan bobot mempertimbangkan jumlah bahan baku dalam satuan *engkel*. Tidak pakai diberi bobot 1 dan 1 *engkel* diberi bobot 2. Begitu juga, untuk tanah pasir sungai, dihitung jumlah *colt* dan diberi bobot tambahan 1. Untuk kaolin, bobotnya ditentukan berdasarkan volume penggunaannya dalam satuan truk. Variasi bobot dimulai dari tidak pakai, yang diberi bobot 1, kemudian 1/4 truk diberi bobot 2, untuk 1/2 truk diberi bobot 3, dan 1 truk diberi bobot 4. Pemberian bobot ini bertujuan untuk memastikan bahwa tidak ada nilai 0 yang dapat mengurangi akurasi dalam perhitungan algoritma K-NN.

Setelah *dataset* dikonversi, dilakukan *splitting data* secara acak dengan perbandingan 70:30, di mana 70% dari total 36 data digunakan sebagai data latih, yaitu 26 data, dan sisanya 30% sebagai data uji atau sebanyak 10 data.

### B. Perhitungan K-NN

Algoritma K-NN mengklasifikasikan hasil *query instance* baru berdasarkan mayoritas kategori K-NN, yang membuatnya termasuk dalam metode *supervised learning* [20]. Implementasi algoritma K-NN dalam penelitian ini melibatkan langkah-langkah seperti pada persamaan (1) sebagai berikut:

#### 1) Memilih jumlah $k$ yang akan digunakan

Perhitungan K-NN dalam penelitian ini pada tahap perhitungan manual menggunakan nilai  $k=3$ . Kemudian setelah dilakukan pemodelan K-NN menggunakan *RapidMiner* nilai  $k=5$  didapatkan hasil yang optimal, sehingga nilai  $k=5$  dipilih sebagai acuan penelitian ini.

#### 2) Hitung jarak data antara data latih dan data uji

Perhitungan jarak dilakukan dengan rumus *Euclidean Distance* yang ditunjukkan pada persamaan (1).

$$d_i = \sqrt{\sum_{i=1}^n (x_1 - x_2)^2} \quad (1)$$

Dalam persamaan (1) tersebut,  $d_i$  mewakili jarak antara dua data,  $x_1$  merupakan *data latih*,  $x_2$  adalah *data uji*,  $i$  adalah variabel yang menunjukkan data, dan  $n$  adalah dimensi data.

#### 3) Pengurutan data dari yang terkecil hingga terbesar dan pemilihan data sesuai jumlah nilai $k$

#### 4) Pengambilan label kelas klasifikasi mayoritas sebagai hasil prediksi

### C. Pemodelan KNN Menggunakan *RapidMiner*

Setelah dilakukan perhitungan secara manual, kemudian dilanjutkan dengan perhitungan menggunakan *RapidMiner* mengoptimalkan hasil klasifikasi menggunakan metode K-NN. Langkah pertama dalam pemodelan K-NN menggunakan *RapidMiner* adalah menyiapkan dua *file excel* terpisah untuk data latih dan uji, kemudian mengimpor keduanya ke dalam *RapidMiner*. Selanjutnya, tentukan atribut yang berfungsi sebagai label atau atribut klasifikasi. Setelah itu, sambungkan data latih dan uji menggunakan *node operator* yang mencakup algoritma K-NN, *Apply Model*, dan *Performance* untuk proses klasifikasi. Dalam *RapidMiner*, operator *Apply Model* digunakan untuk menerapkan model pada data baru untuk membuat prediksi, sementara operator *Performance* digunakan untuk mengevaluasi kinerja model klasifikasi dengan berbagai metrik, dalam penelitian ini metrik yang digunakan adalah akurasi.

Tujuan pemodelan K-NN menggunakan *RapidMiner* adalah untuk menilai akurasi dari setiap nilai  $k$ , yang dalam penelitian ini menggunakan nilai  $k$  1, 3, 5, 7, dan 9. Nilai akurasi tertinggi didapatkan pada nilai  $k$  1, 3, 5 sebesar 90% kemudian dalam penelitian ini nilai  $k=5$  dipilih sebagai nilai tengah  $k$  yang memberikan akurasi tertinggi selama pengujian [16].

### D. Implementasi Sistem

Tahap implementasi adalah penulisan kode program berdasarkan desain hasil perancangan [21]. Sistem dikembangkan dengan menggunakan bahasa pemrograman PHP dan basis data MySQL. Dimulai dari menerjemahkan analisis kebutuhan menjadi desain sistem yang bisa diimplementasikan. Perancangan sistem mencakup alur penyimpanan *dataset*, pengolahan awal data untuk memenuhi syarat perhitungan K-NN, struktur tabel basis data, dan tampilan hasil prediksi.

### E. Pengujian

Dalam penelitian ini, terdapat dua metode pengujian. Pada metode pertama, dilakukan perhitungan akurasi antara klasifikasi data hasil wawancara, hasil pemodelan K-NN menggunakan *RapidMiner*, dan hasil prediksi sistem. Pengujian akurasi dilakukan untuk mengetahui tingkat keberhasilan algoritma K-NN dalam memprediksi kualitas genteng mantili [12]. Pengujian dapat dilakukan dengan menggunakan persamaan (2).

$$akurasi = \frac{\text{nilai benar}}{\text{jumlah data}} \times 100\% \quad (2)$$

Di mana nilai benar dalam konteks ini merujuk pada jumlah data yang memberikan hasil yang sesuai atau benar dalam pengujian. Sedangkan jumlah data merupakan jumlah data yang diuji [12]. Sedangkan pada metode kedua, dilakukan pengujian sistem menggunakan empat data aktual di tempat penelitian. Empat data yang digunakan untuk pengujian hasil prediksi sistem merupakan data dengan klasifikasi tidak bagus pilihan dari narasumber. Hasil analisis dari metode kedua menyoroti sejauh mana prediksi sistem cocok dengan data aktual, memberikan wawasan yang lebih dalam tentang keandalan sistem.

## III. HASIL DAN PEMBAHASAN

### A. Pengumpulan dan Pengolahan Data

#### 1) Data hasil wawancara ditampilkan pada Tabel 1.

TABEL I  
DATA HASIL WAWANCARA

No	Tanah Liat Lempung	Tanah Liat Hitam	Tanah Liat Merah	Tanah Liat Padas	Pasir Sungai atau Ladu	Kaolin	Klasifikasi
1	1 engkel	2 engkel	5 engkel	3 engkel	1 colt	1 truk	tidak bagus
2	1 engkel	5 engkel	2 engkel	3 engkel	1 colt	1/2 truk	bagus
3	1 engkel	2 engkel	6 engkel	4 engkel	1 colt	1/4 truk	tidak bagus
4	1 engkel	tidak pakai	tidak pakai	1 engkel	2 colt	1/4 truk	tidak bagus
5	1 engkel	1 engkel	1 engkel	2 engkel	2 colt	tidak pakai	tidak bagus
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
35	3 engkel	tidak pakai	1 engkel	1 engkel	4 colt	1/4 truk	tidak bagus
36	4 engkel	3 engkel	2 engkel	tidak pakai	1 colt	tidak pakai	tidak bagus

#### 2) Tabel kriteria dan sub-kriteria: terdapat 6 kriteria dan sub-kriteria beserta nilai pembobotan sebagaimana pada Tabel 2.

TABEL II  
KRITERIA DAN SUB-KRITERIA

Kriteria	Nama Kriteria	Sub-Kriteria	Bobot
T1	Tanah Liat Lempung	tidak pakai	1
		1 engkel	2
		2 engkel	3
T2	Tanah Liat Hitam	tidak pakai	1
		1 engkel	2
		2 engkel	3
T3	Tanah Liat Merah	tidak pakai	1
		1 engkel	2
		2 engkel	3
T4	Tanah Liat Padas	tidak pakai	1
		1 engkel	2
		2 engkel	3
T5	Tanah Pasir Sungai	tidak pakai	1
		1 colt	2
		2 colt	3
T6	Kaolin	tidak pakai	1
		1/4 truk	2
		1/2 truk	3
		1 truk	4

### 3) Melakukan pembobotan data

Data yang didapatkan setelah wawancara dilakukan konversi dengan kriteria dan sub-kriteria, hasil dapat dilihat pada Tabel 2. Untuk ketentuan konversi pembobotan secara rinci dijelaskan pada Tabel 1. *Dataset* hasil konversi bobot terdiri dari 36 data, kemudian dipisahkan dengan *split validation* 70% : 30% menjadi 26 data latih yang terdapat dalam Tabel 3 dan 10 data uji yang terdapat dalam Tabel 4. Data latih dan data uji inilah yang akan dipergunakan dalam perhitungan dengan menggunakan algoritma K-NN.

TABEL III  
 DATASET HASIL KONVERSI BOBOT

No	T1	T2	T3	T4	T5	T6	Klasifikasi
1	2	3	6	4	2	2	tidak bagus
2	2	6	3	4	2	2	bagus
3	2	3	7	5	2	2	tidak bagus
4	2	1	1	2	3	2	tidak bagus
5	2	2	2	3	3	1	tidak bagus
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
35	4	1	2	2	5	2	tidak bagus
36	5	4	3	1	2	1	tidak bagus

TABEL IV  
 DATA LATIH

No	T1	T2	T3	T4	T5	T6	Klasifikasi
1	2	3	6	4	2	2	tidak bagus
2	2	6	3	4	2	2	bagus
3	2	3	7	5	2	2	tidak bagus
4	2	1	1	2	3	2	tidak bagus
5	2	2	2	3	3	1	tidak bagus
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
25	1	5	2	3	2	2	bagus
26	1	2	3	2	4	1	tidak bagus

TABEL V  
 DATA UJI

No	T1	T2	T3	T4	T5	T6	Klasifikasi
1	2	4	2	4	1	2	bagus
2	2	1	1	3	2	1	tidak bagus
3	3	2	6	5	2	2	tidak bagus
4	3	3	2	5	5	2	tidak bagus
5	3	3	2	2	1	2	bagus
6	3	5	2	4	2	2	bagus
7	4	3	1	3	2	2	bagus
8	4	6	3	5	3	2	bagus
9	4	1	2	2	5	2	tidak bagus
10	5	4	3	1	2	1	tidak bagus

### B. Perhitungan K-NN

#### 1) Menentukan nilai $k=3$

#### 2) Mengambil satu data uji yaitu nomor 8 kemudian dilakukan perhitungan jarak terhadap semua data latih

Untuk menghitung jarak antara data latih dan data uji dapat menggunakan persamaan (1). Satu data uji dilakukan perhitungan jarak terhadap keseluruhan data latih. Berikut adalah contoh perhitungan jarak antara data uji nomor 8 dengan data latih nomor 1 dapat dilihat pada persamaan (3).

$$d_i = \sqrt{\sum_{i=1}^n (x_1 - x_2)^2}$$

$$d_1 = 4.898979486 \tag{3}$$

3) Mengurutkan hasil hitung jarak dimulai dari yang terkecil kemudian diambil data sesuai jumlah  $k=3$ , hasil dapat dilihat pada Tabel 6.

TABEL VI  
 HASIL PENGURUTAN DATA

No	T1	T2	T3	T4	T5	T6	Nilai ED	Sorting	Klasifikasi
22	2	3	6	4	2	2	1,414	1	bagus
2	2	6	3	4	2	2	2,449	2	bagus
20	5	4	3	1	2	1	3	3	bagus

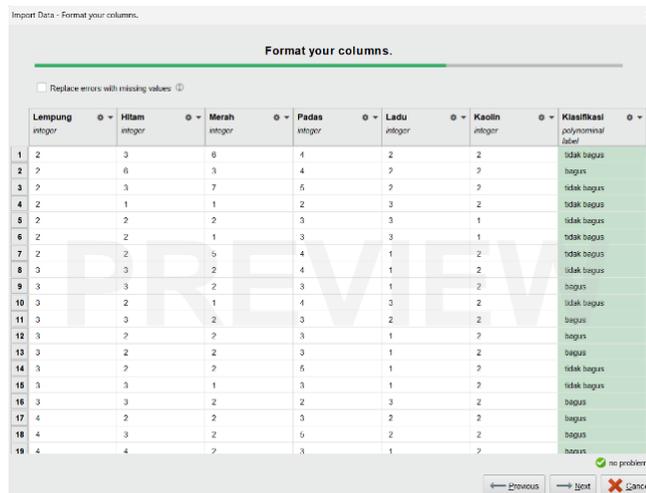
4) Pengambilan label kelas klasifikasi mayoritas sebagai hasil prediksi

Setelah data diurutkan dari terkecil diambil hasil klasifikasi paling banyak muncul, maka dari pengujian data uji nomor 8 didapatkan hasil prediksi sebagai bagus dan prediksi tersebut sesuai dengan data asli.

### C. Pemodelan K-NN Menggunakan RapidMiner

RapidMiner adalah perangkat lunak yang digunakan untuk analisis data melalui algoritma penambangan data [22]. Integrasi mencakup keseluruhan proses ilmu data, mulai dari persiapan data hingga pemanfaatan model prediktif [23]. RapidMiner mempermudah pengguna dalam menganalisis data besar dengan menggunakan operator yang dapat mengubah data. Data yang telah terhubung ke *node* pada operator tersebut kemudian dapat dihubungkan ke *node* hasil untuk melihat hasil akhirnya [22].

Langkah awal yang perlu dilakukan untuk memulai pemodelan K-NN menggunakan RapidMiner adalah menyiapkan dua *file excel* yang berbeda, satu berisi data latih dan yang lainnya berisi data uji. Setelah *file excel* tersedia, selanjutnya mengimpornya ke RapidMiner. Kemudian atur atribut yang akan dijadikan label, yaitu atribut klasifikasi. Secara rinci, proses pemilihan atribut data sebagai label ditunjukkan pada Gambar 1.



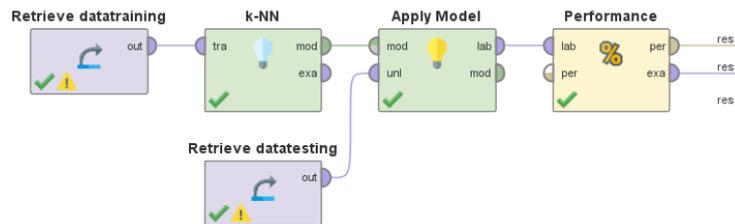
Gambar 1. Menentukan Atribut Label Dataset

Untuk memulai analisis ditunjukkan pada Gambar 2, Hubungkan data latih dan data uji dengan *node* operator yang terdiri dari algoritma K-NN, *Apply Model*, dan *Performance* untuk klasifikasi. Setelah itu, modifikasi nilai  $k$  sesuai kebutuhan. Nilai  $k$  yang digunakan dalam penelitian ini adalah 1, 3, 5, 7, dan 9. Informasi mengenai nilai akurasi yang diperoleh akan ditampilkan pada layar.

Setelah pembentukan model pola K-NN, penelitian ini menguji setiap nilai  $k$ , mulai dari angka 1, 3, 5, 7, dan 9. Tujuannya adalah untuk mengukur tingkat akurasi yang dihasilkan oleh setiap nilai  $k$  terhadap klasifikasi kualitas genteng mantili yang diteliti. Rumus perhitungan akurasi dapat dilihat pada persamaan (2). Untuk hasil nilai akurasi secara rinci dijelaskan pada Tabel 7.

TABEL VII  
 HASIL AKURASI RAPIDMINER

Nilai K	Jumlah Data	Jumlah Benar	Jumlah Salah	Akurasi
1	10	9	1	90%
3	10	9	1	90%
5	10	9	1	90%
7	10	8	2	80%
9	10	8	2	80%



Gambar. 2. Pemodelan K-NN Menggunakan *RapidMiner*

Hasil pengujian algoritma K-NN menunjukkan akurasi yang sangat baik, dengan tingkat akurasi 90% dan 80% [24]. Nilai  $k=5$  memberikan akurasi tinggi saat pengujian [16]. Akurasi K-NN dipengaruhi pemilihan nilai  $k$ , apabila nilai  $k$  kecil, akan sensitif terhadap *noise*. Namun, jika terlalu besar dapat menyebabkan bias pada model [20]. Hasil akurasi dari penentuan nilai  $k$  akan menjadi acuan dalam implementasi sistem.

Pemanfaatan algoritma K-NN dalam prediksi kualitas genteng mantili berdasarkan komposisi bahan baku menghadirkan manfaat seperti kesederhanaan dan kemampuan dalam memperkirakan hasil [10]. Namun, Ketika terjadi kelangkaan data bahan baku atau situasi produksi yang tidak biasa, penerapan K-NN untuk memprediksi kualitas genteng menjadi sulit karena model memerlukan data yang cukup untuk membuat prediksi yang akurat. *RapidMiner*, alat untuk pemodelan K-NN, dapat memberikan prediksi akurat untuk berbagai skenario, termasuk memprediksi penjualan motor terlaris [25]. Terlepas dari efektivitasnya, *RapidMiner* mungkin memiliki keterbatasan dibandingkan dengan alat lain, seperti yang terlihat dalam sebuah penelitian di mana WEKA mengungguli *RapidMiner* dalam akurasi untuk memprediksi ketepatan waktu kelulusan di universitas [26]. Untuk mengatasi tantangan ini, sebelum dilakukan perhitungan dengan model sebaiknya memperkaya jumlah *dataset* yang digunakan kemudian melakukan tahap *preprocessing data* secara teliti seperti penanganan *outlier*, normalisasi data, dan penambahan atribut baru yang relevan. Penelitian telah menunjukkan bahwa menganalisis sentimen terkait perpindahan Ibu Kota dapat meningkatkan akurasi K-NN menjadi 88,12% [13].

#### D. Implementasi Sistem

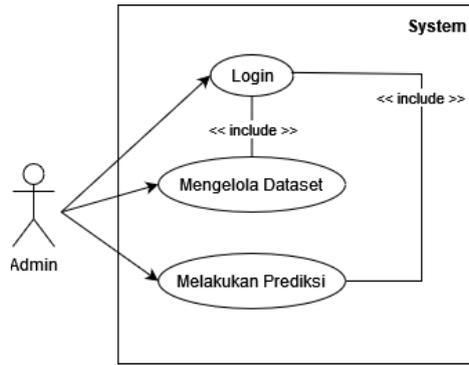
Bagan alur sistem yang menggambarkan proses secara umum dapat dilihat pada Gambar 3. Alur sistem dimulai dengan memasukkan data bahan baku, dilanjutkan dengan melakukan perhitungan menggunakan algoritma K-NN, dan terakhir mengambil kesimpulan dari hasil perhitungan sebagai prediksi akhir [20].



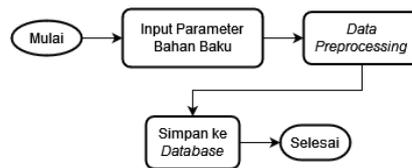
Gambar. 3. Bagan Alur Sistem

Berdasarkan alur sistem, hasil analisis kebutuhan sistem tergambar dalam diagram *use case* yang ditampilkan dalam Gambar 4. Diagram *use case* mengilustrasikan keterkaitan antara aktor dan sistem dalam interaksi mereka [8]. Dalam diagram *use case* menunjukkan semua modul dalam sistem ini mengharuskan login terlebih dahulu untuk dapat digunakan [27]. Pemodelan proses sistem hanya melibatkan satu aktor, yaitu admin. Kebutuhan sistem ini mencakup dua hal utama, yaitu mengelola *dataset* dan melakukan prediksi kualitas genteng mantili.

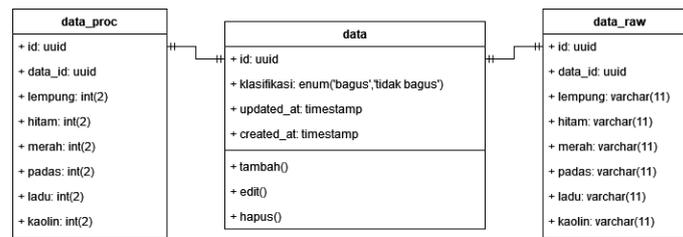
Data bahan baku yang didapatkan bertipe tidak numerik, maka diperlukan pengolahan data awal agar dapat diubah menjadi data numerik [19]. Data yang dimasukkan oleh pengguna akan diubah terlebih dahulu sesuai dengan ketentuan konversi pembobotan dapat dilihat pada Tabel 1. Sebelum kemudian disimpan ke dalam *database*. Proses alur pengolahan data awal dalam sistem dijelaskan secara rinci dalam Gambar 5.



Gambar. 4. Use Case Diagram Sistem



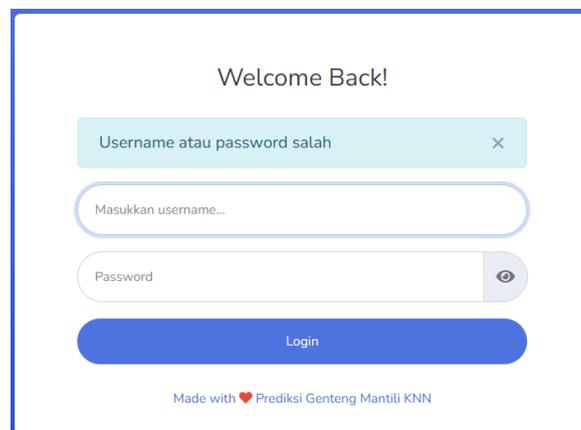
Gambar. 5. Alur Menyimpan Dataset Pada Sistem



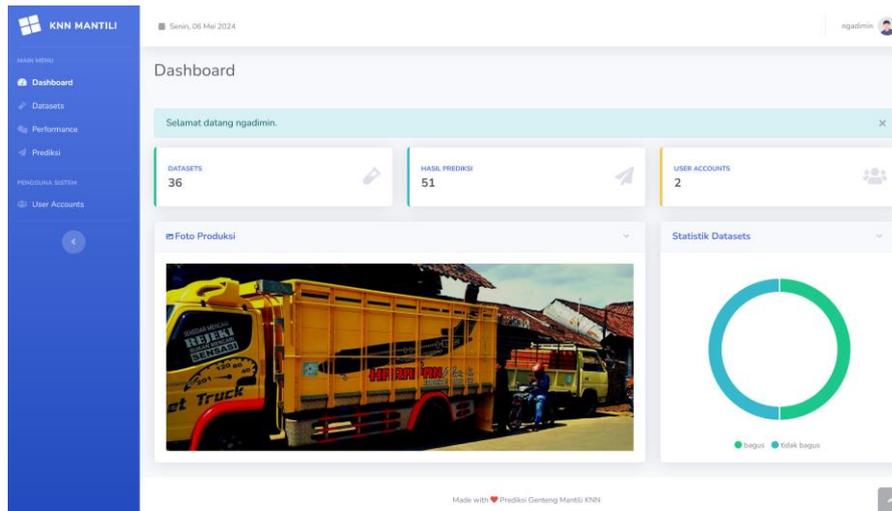
Gambar. 6. Class Diagram Sistem

Untuk memenuhi kebutuhan pengolahan data awal, desain *database* membutuhkan kemampuan untuk menyimpan data dalam dua kondisi, yaitu sebelum dan setelah diproses. Oleh karena itu, dibuat tiga tabel dengan nama “data” untuk menyimpan informasi data secara umum, “data\_proc” untuk menyimpan data yang telah diolah, dan “data\_raw” untuk menyimpan data mentah. Tabel data berelasi secara *one to one* terhadap tabel data\_proc. Tabel data berelasi secara *one to one* terhadap tabel data\_raw. Relasi dari ketiga tabel ini dalam penggunaannya dapat disesuaikan dengan kebutuhan. Secara rinci desain *database* ditunjukkan pada Gambar 6.

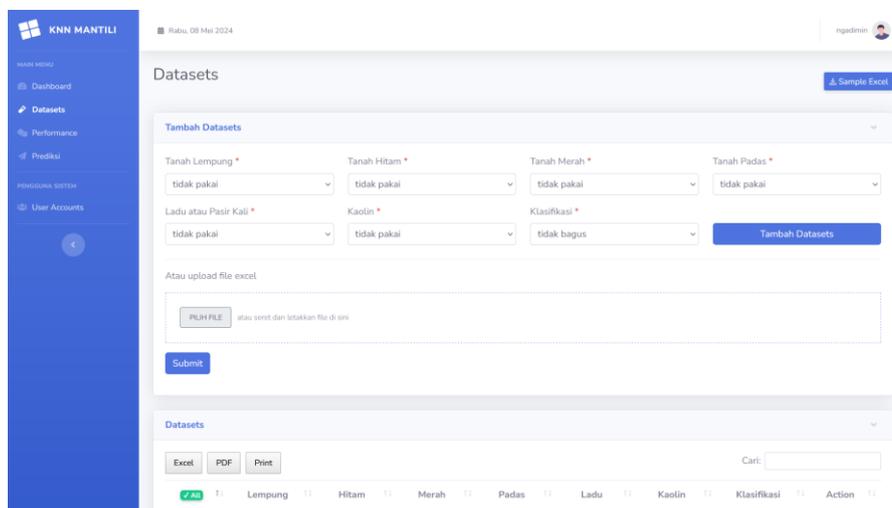
Hasil implementasi tampilan yang pertama adalah halaman login. Halaman login akan mengarahkan ke *dashboard* apabila skema login tervalidasi oleh sistem. Pengguna wajib mengisi *username* dan *password*. Ketika terjadi kesalahan saat login. Sistem akan memberitahukan informasi kesalahan yang terjadi apabila *username* atau *password* yang dimasukkan tidak valid. Secara rinci tampilan halaman login dapat dilihat pada Gambar 7.



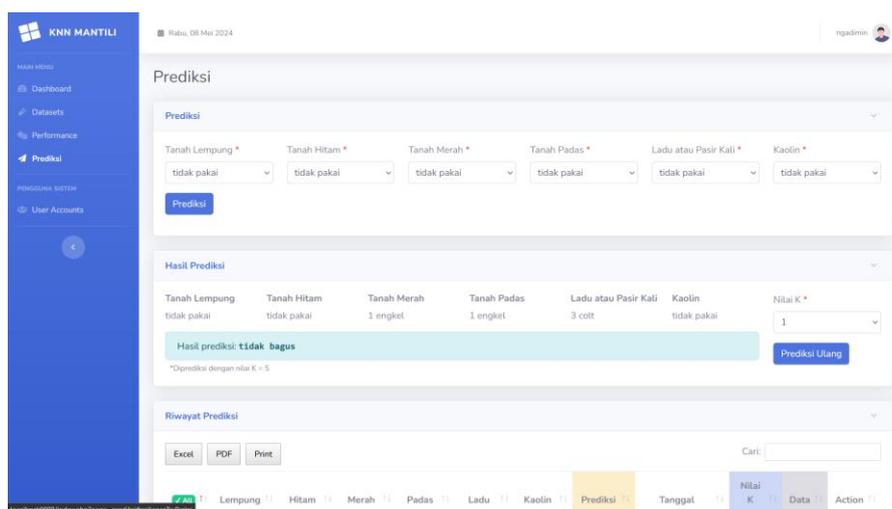
Gambar. 7. Implementasi Halaman Login Sistem



Gambar. 8. Implementasi Halaman *Dashboard*



Gambar. 9. Implementasi Halaman Kelola *Dataset*



Gambar. 10. Implementasi Halaman *Prediksi*

Gambar 8 memperlihatkan implementasi halaman *dashboard* setelah pengguna berhasil login. Berisi informasi jumlah *dataset* yang tersimpan, statistik dari *dataset* berupa informasi total data dengan klasifikasi “bagus” dan “tidak bagus”, jumlah hasil prediksi, dan album foto produksi dari pengrajin genteng Bapak Sigit. Untuk bagian *sidebar* terdapat menu mengelola *dataset* dan menu prediksi.

Gambar 9 memperlihatkan implementasi halaman kelola *dataset*. Terdapat *form* input *dataset* dan tabel yang berisikan keseluruhan *dataset* yang tersimpan. Dalam halaman ini pengguna juga bisa melakukan tambah, edit, dan hapus *dataset*.

Gambar 10 memperlihatkan implementasi halaman prediksi. Terdapat *form* inputan parameter bahan baku untuk melakukan prediksi kualitas genteng mantili. Apabila *form* prediksi dikirim, sistem akan melakukan perhitungan K-NN dengan nilai *default*  $k=5$ . Informasi yang terdapat pada tampilan hasil prediksi diantaranya adalah parameter bahan baku, hasil prediksi, nilai  $k$  untuk melakukan prediksi, dan disediakan *form* untuk melakukan prediksi ulang menggunakan nilai  $k$  yang berbeda.

### E. Pengujian

Dalam penelitian ini, ada dua metode pengujian. Metode pertama memeriksa klasifikasi data hasil wawancara, hasil pemodelan K-NN menggunakan *RapidMiner* dan hasil prediksi sistem. Metode kedua menguji sistem dengan 4 data aktual di lapangan untuk mengevaluasi akurasi hasil prediksi. Secara rinci hasil pengujian pertama dapat dilihat pada Tabel 8.

TABEL VIII  
 HASIL PREDIKSI DATA UJI

Data Nomor	Data Asli	Prediksi Sistem	RapidMiner	Keterangan
1	bagus	bagus	bagus	sama
2	tidak bagus	tidak bagus	tidak bagus	sama
3	tidak bagus	bagus	bagus	sama
4	tidak bagus	bagus	bagus	sama
5	bagus	bagus	tidak bagus	tidak sama
6	bagus	tidak bagus	tidak bagus	sama
7	bagus	bagus	bagus	sama
8	bagus	bagus	bagus	sama
9	tidak bagus	tidak bagus	tidak bagus	sama
10	tidak bagus	bagus	bagus	tidak sama

Berdasarkan hasil prediksi pada Tabel 8, dari 10 data yang diuji, terdapat 2 data yang hasil prediksinya tidak sama. Dari 10 data tersebut, 8 di antaranya diprediksi secara benar. Dengan menggunakan persamaan (2) untuk keseluruhan data, didapatkan akurasi sebesar 80%. Data yang memiliki hasil prediksi berbeda atau salah adalah data nomor 5 dan data nomor 10. Dengan demikian, hasil prediksi dengan aplikasi menghasilkan *error rate* sebesar 20%.

Hasil pengujian menunjukkan bahwa metode K-NN dapat memprediksi kualitas genteng mantili dengan baik, yaitu sekitar 80%. Ini artinya, sistem bisa memprediksi apakah genteng akan bagus atau tidak berdasarkan bahan bakunya. Akurasi yang memuaskan ini memberikan dasar kuat bagi pengguna sistem untuk menggunakan prediksi model dalam mengoptimalkan proses produksi genteng, mengurangi pemborosan sumber daya, dan meningkatkan kualitas produk secara keseluruhan. Dengan demikian, keberhasilan model K-NN dalam menghasilkan prediksi yang akurat menyoroti nilai penting metode ini dalam konteks penggunaan praktis dalam industri genteng.

TABEL IX  
 HASIL PREDIKSI DATA AKTUAL LAPANGAN

No	Tanah Liat Lempung	Tanah Liat Hitam	Tanah Liat Merah	Tanah Liat Padas	Pasir Sungai atau Ladu	Kaolin	Prediksi Sistem	Keterangan
1	tidak pakai	1 engkel	4 engkel	1 engkel	tidak pakai	1/4 truk	tidak bagus	sesuai
2	tidak pakai	tidak pakai	1 engkel	3 engkel	1 colt	tidak pakai	tidak bagus	sesuai
3	4 engkel	4 engkel	tidak pakai	1 engkel	1 colt	tidak pakai	tidak bagus	sesuai
4	tidak pakai	tidak pakai	1 engkel	1 engkel	3 colt	tidak pakai	tidak bagus	sesuai

Pada uji coba kedua yang menggunakan nilai  $k=5$ , sistem dimasukkan data aktual yang diperoleh di lapangan [12]. Data yang digunakan dalam uji coba ini memiliki klasifikasi "tidak bagus". Uji coba bertujuan untuk menguji kemampuan sistem dalam memberikan akurasi yang baik serta mencegah terjadinya prediksi yang salah yang dapat merugikan pengrajin genteng mantili. Hasil prediksi yang didapatkan sistem dapat memprediksi data aktual dengan sesuai. Detail hasil prediksi dapat dilihat pada Tabel 9.

Dari data yang terdapat pada Tabel 8 dan Tabel 9, dapat disimpulkan bahwa sistem yang dibangun telah menghasilkan akurasi yang bagus. Hal ini menunjukkan bahwa algoritma K-NN dapat diterapkan untuk memprediksi kualitas genteng mantili berdasarkan komposisi bahan baku.

Perbandingan hasil penelitian serupa yang dilakukan sebelumnya yaitu dalam klasifikasi data penerima zakat menggunakan 6 fitur meliputi umur, tanggungan, pekerjaan, jumlah usulan, penghasilan, dan kondisi rumah. Rasio data uji dan latih 70% : 30% serta nilai  $k=5$ , mencapai akurasi 86% [28]. Perbedaan terletak pada pendekatan *preprocessing data*, di mana metode *min-max* diterapkan untuk menormalisasi nilai-nilai fitur ke dalam rentang 0 hingga 1 dengan mengurangi nilai minimum dan membaginya dengan rentang nilai fitur.

Meskipun algoritma K-NN telah berhasil menghasilkan akurasi yang cukup tinggi dalam memprediksi kualitas genteng mantili berdasarkan komposisi bahan baku, penelitian selanjutnya direkomendasikan untuk mengeksplorasi dengan menerapkan pendekatan metodologi yang berbeda agar memperoleh hasil yang lebih beragam dan relevan untuk perbandingan dengan penelitian ini.

#### IV. KESIMPULAN

Hasil penelitian ini menyimpulkan bahwa algoritma K-NN dengan *Euclidean Distance* dapat diterapkan untuk memprediksi kualitas genteng mantili berdasarkan komposisi bahan baku. Hasil pengujian metode akurasi dengan menggunakan 26 data latih dan 10 data uji menunjukkan bahwa tingkat akurasi mencapai 80% dengan nilai *error* sebesar 20%. Akurasi diperoleh dari perbandingan antara klasifikasi data hasil wawancara, hasil pemodelan K-NN menggunakan *RapidMiner*, dan hasil prediksi sistem. Pengujian kedua dengan memasukkan data aktual di lapangan menunjukkan hasil yang sesuai dengan situasi di lapangan. Adapun saran untuk penelitian yang ingin mengembangkan lebih lanjut adalah pada penelitian selanjutnya diharapkan dapat ditambahkan fitur yang dapat disesuaikan dengan komposisi bahan baku yang digunakan oleh pengrajin genteng.

#### DAFTAR PUSTAKA

- [1] D. Nirmalasari, I. H. Lubis, H. E. Kusuma, and M. D. Koerniawan, "Preferensi Penggunaan Material pada Atap Rumah Tinggal," *Tesa Arsitektur*, vol. 18, no. 1, pp. 1–9, 2020.
- [2] D. Y. Reindrawati et al., *Analisis Teori dan Pengaplikasian Model Bisnis Canvas pada Usaha Kecil Genteng Tanah Liat Tradisional di Kediri, Jawa Timur*. Lakeiskha, 2020.
- [3] I. G. A. A. Wulandari and N. M. I. Piliandani, "Pemberdayaan UMKM Pengrajin Genteng Tanah Liat Di Desa Pejaten, Kediri-Tabanan, Bali," *Akuntansi dan Humaniora: Jurnal Pengabdian Masyarakat*, vol. 1, no. 2, pp. 78–81, 2022.
- [4] A. Khairunisa, Aviasti, and L. Nurwandi, "Usulan Perbaikan Proses Pembuatan Genteng dengan Menggunakan Metode Taguchi pada Home Industri Mahkota," *Bandung Conference Series: Industrial Engineering Science*, vol. 3, no. 1, pp. 268–275, Jan. 2023, doi: 10.29313/bcsies.v3i1.6614.
- [5] A. Z. Rohmah and L. Ayundasari, "Pengaruh industri genteng tanah liat terhadap perubahan sosial ekonomi masyarakat Desa Kamulan Tahun 2001-2021," *Jurnal Integrasi dan Harmoni Inovatif Ilmu-Ilmu Sosial*, vol. 2, no. 3, pp. 223–232, 2022.
- [6] B. A. Prastyo, "Perancangan Sistem Penggorengan Pada Mesin Pembuat Keripik Serbaguna dengan Metode Deep Frying," *Prosiding SEMNAS INOTEK (Seminar Nasional Inovasi Teknologi)*, vol. 4, no. 3, pp. 285–290, Aug. 2020, doi: 10.29407/inotek.v4i3.101.
- [7] M. Faris Raharja and M. E. Yuliana, "Teknologi Informasi Sebagai Sarana Komunikasi Penjual kepada Pembeli di Platform E-Commerce," *Jurnal Ekonomi Teknologi & Bisnis (JETBIS)*, vol. 1, no. 3, 2022, [Online]. Available: <https://jetbis.al-makkipublisher.com/index.php/al/index>
- [8] V. Atina and D. Hartanti, "Knowledge Based Recommendation Modeling for Clothing Product Selection Recommendation System," *Jurnal Teknik Informatika (Jutif)*, vol. 3, no. 5, pp. 1407–1413, Oct. 2022, doi: 10.20884/1.jutif.2022.3.5.584.
- [9] M. Mufarrohah, T. Yulianto, and F. Faisol, "Penentuan Jenis Tanah untuk Menghasilkan Genteng Berkualitas Menggunakan Fuzzy ELECTRE," *Zeta - Math Journal*, vol. 5, no. 1, pp. 32–36, Nov. 2020, doi: 10.31102/zeta.2020.5.1.32-36.
- [10] R. N. Sukmana, A. Abdurrahman, and Y. Wicaksono, "Implementasi K-Nearest Neighbor untuk Menentukan Prediksi Penjualan: (Studi Kasus: PT MaksIPlus Utama Indonesia)," *Jurnal Teknologi Informasi dan Komunikasi*, vol. 9, no. 2, pp. 31–37, Dec. 2020, doi: 10.58761/jurtikstmikbandung.v9i2.123.
- [11] K. Supriyanto, D. Whizkid Aziiz, H. N. Gian, and D. Hartanti, "Sistem Pendukung Keputusan Kualitas Genteng di Desa Wirun Menggunakan Metode Multi Attribute Utility Theory (MAUT)," in *Prosiding Seminar Nasional Teknologi Informasi dan Bisnis*, 2022, pp. 338–342.
- [12] A. Yudhana, S. Sunardi, and A. J. S. Hartanta, "Algoritma K-NN dengan Euclidean Distance untuk Prediksi Hasil Penggajian Kayu Sengon," *Transmisi*, vol. 22, no. 4, pp. 123–129, Nov. 2020, doi: 10.14710/transmisi.22.4.123-129.
- [13] S. D. Prasetyo, S. S. Hilabi, and F. Nurapriani, "Analisis Sentimen Relokasi Ibukota Nusantara Menggunakan Algoritma Naïve Bayes dan KNN," *Jurnal KomtekInfo*, pp. 1–7, Jan. 2023, doi: 10.35134/komtekinfo.v10i1.330.
- [14] Sahar, "Analisis Perbandingan Metode K-Nearest Neighbor dan Naïve Bayes Clasiffier Pada Dataset Penyakit Jantung," *Indonesian Journal of Data and Science*, vol. 1, no. 3, pp. 79–86, 2020, doi: <https://doi.org/10.33096/ijodas.v1i3.20>.
- [15] Y. Reswan, R. Toyib, H. Witriyono, and A. Anggraini, "Klasifikasi Tingkat Kematangan Buah Nanas Berdasarkan Fitur Warna Menggunakan Metode K-Nearest Neighbor (KNN)," *JURNAL MEDIA INFOTAMA*, vol. 20, no. 1, pp. 280–287, Apr. 2024, doi: 10.37676/jmi.v20i1.5689.
- [16] C. Anisa and Andri, "Penerapan Algoritma K-Nearest Neighbor untuk Prediksi Penjualan Obat pada Apotek Kimia Farma Atmo Palembang," in *Bina Darma Conference on Computer Science (BDCCS)*, 2020, pp. 199–208.
- [17] Y. Kurniawan and T. Barokah, "Klasifikasi Penentuan Pengajuan Kartu Kredit Menggunakan K-Nearest Neighbor," *Jurnal Ilmiah Matrik*, vol. 22, no. 1, pp. 73–82, Mar. 2020, doi: 10.33557/jurnalmatrrik.v22i1.843.
- [18] S. Sukanto, Y. Adriyani, and R. Aulia, "Prediksi Kelompok UKT Mahasiswa Menggunakan Algoritma K-Nearest Neighbor," *JUITA: Jurnal Informatika*, vol. 8, no. 1, pp. 121–130, 2020.
- [19] K. M. Rajabi, W. Witanti, and R. Yuniarti, "Penerapan Algoritma K-Nearest Neighbor (KNN) Dengan Fitur Relief-F Dalam Penentuan Status Stunting," *Innovative: Journal of Social Science Research*, vol. 3, no. 4, pp. 3555–3568, 2023, doi: 10.31004/innovative.v3i4.3885.
- [20] S. R. Cholil, T. Handayani, R. Prathivi, and T. Ardianita, "Implementasi Algoritma Klasifikasi K-Nearest Neighbor (KNN) Untuk Klasifikasi Seleksi Penerima Beasiswa," *IJCIT (Indonesian Journal on Computer and Information Technology)*, vol. 6, no. 2, pp. 118–127, 2021.
- [21] V. A. Lestari, A. Yuli Ananta, and P. Basudewa, "Sistem Informasi Prediksi Persediaan Obat Di Apotek Naylun Farma Menggunakan Holt-Winters," *JIP (Jurnal Informatika Polinema)*, vol. 9, no. 2, pp. 229–236, 2023.

- [22] L. Widyawati and V. Lusiana, "Penerapan K-Means Clustering Untuk Mengelompokkan Data Transaksi Penjualan (Studi Kasus pada Wijaya Hijab)," *Jutisi: Jurnal Ilmiah Teknik Informatika dan Sistem Informasi*, vol. 12, no. 3, 2023.
- [23] V. R. Prasetyo, H. Lazuardi, A. A. Mulyono, and C. Lauw, "Penerapan Aplikasi RapidMiner Untuk Prediksi Nilai Tukar Rupiah Terhadap US Dollar Dengan Metode Regresi Linier," *Jurnal Nasional Teknologi dan Sistem Informasi (TEKNOSI)*, vol. 7, no. 1, pp. 8–17, 2021.
- [24] 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.
- [25] Rismala, I. Ali, and A. Rizki Rinaldi, "Penerapan Metode K-Nearest Neighbor untuk Prediksi Penjualan Sepeda Motor Terlaris," *Jurnal Mahasiswa Teknik Informatika*, vol. 7, no. 1, pp. 585–590, 2023.
- [26] D. Damayanti, "Perbandingan Akurasi Software Rapidminer dan Weka Menggunakan Algoritma K-Nearest Neighbor (K-NN)," *Jurnal Syntax Admiration*, vol. 2, no. 6, pp. 994–1006, Jun. 2021, doi: 10.46799/jsa.v2i6.247.
- [27] D. Christy Pratiwi, V. Atina, and J. Maulindar, "Job-Position Recommender System Using Knowledge Based Recommendation Method at Atmi Polytechnic Surakarta," *Jurnal Teknik Informatika (JUTIF)*, vol. 5, no. 1, pp. 153–161, 2024, doi: 10.52436/1.jutif.2024.5.1.1258.
- [28] A. Hernandez, S. K. Gusti, F. Syafria, L. Handayani, and S. Ramadhani, "Klasifikasi Data Penerimaan Zakat dengan Algoritma K-Nearest Neighbor," *KLIK: Kajian Ilmiah Informatika dan Komputer*, vol. 4, no. 3, pp. 1632–1640, 2023.