

# Vaname Shrimp Harvest Prediction System Using The K-Nearest Neighbors (KNN) Method

Ahmad Ubaidillah<sup>1\*</sup>, Retno Wardhani<sup>2</sup>, Ayu Ismi Hanifah<sup>3</sup>

<sup>1,2,3</sup> Program Studi Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Islam Lamongan, Jawa Timur, Indonesia  
[ahmadubaidillah201@gmail.com](mailto:ahmadubaidillah201@gmail.com)<sup>1</sup>, [retno.teknik@unisla.ac.id](mailto:retno.teknik@unisla.ac.id)<sup>2</sup>, [ayuismihanifah@gmail.com](mailto:ayuismihanifah@gmail.com)<sup>3</sup>

## Abstrak

Udang vaname merupakan salah satu komoditas budidaya budidaya yang bernilai ekonomi tinggi dan banyak dibudidayakan di wilayah pesisir Indonesia, termasuk di Desa Pendowolimo, Kecamatan Karangbinayun, Kabupaten Lamongan. Namun, produktivitas panen seringkali tidak stabil karena faktor-faktor seperti perubahan iklim, kualitas air, dan terbatasnya penggunaan teknologi prediktif. Penelitian ini bertujuan untuk merancang dan membangun sistem prediksi panen udang vaname berbasis web menggunakan metode K-Nearest Neighbors (KNN). Metode KNN dipilih karena kemampuannya melakukan klasifikasi dan prediksi berdasarkan kedekatan data historis. Kumpulan data yang digunakan terdiri dari 150 catatan panen dengan variabel antara lain periode panen, luas tambak, jumlah tebar ("rean"), durasi tumbuh, dan jumlah pakan. Perhitungan dilakukan dengan menggunakan jarak Euclidean untuk menemukan tetangga terdekat K tersebut, dan hasil panen diprediksi dari nilai rata-rata tetangga tersebut. Untuk mengukur akurasi sistem, digunakan Mean Absolute Percentage Error (MAPE), menghasilkan MAPE sebesar 9,6%, yang dikategorikan sebagai prediksi yang sangat baik. Sistem ini diimplementasikan pada platform web dengan fitur login, manajemen data, dan prediksi real-time otomatis. Hasil penelitian ini menunjukkan bahwa menggunakan metode KNN untuk memprediksi panen udang vaname mampu membantu petambak membuat keputusan budidaya yang lebih terencana dan berbasis data. Oleh karena itu, sistem tersebut diharapkan dapat meningkatkan efisiensi, produktivitas, dan keberhasilan panen di masa depan.

**Kata kunci:** K-Tetangga Terdekat, Udang Vaname, Prediksi Panen, MAPE

## 1. Pendahuluan

Udang kaki putih merupakan jenis udang yang banyak dibudidayakan di Indonesia karena popularitasnya di kalangan petambak dan nilai ekonominya yang tinggi. Udang ini juga banyak digunakan dalam berbagai olahan makanan. Berasal dari daerah subtropis di Pantai Barat Amerika Serikat, udang kaki putih dapat hidup di laut dan air tawar. Mereka panjang, ramping, dan kompak, dengan kandungan protein tinggi dan harga yang relatif terjangkau. [1].

Habitat asli udang kaki putih berada di perairan Pasifik, tetapi spesies ini dapat dibudidayakan dengan baik di Indonesia karena kondisi iklim dan lingkungan yang sesuai [2]. Keunggulan udang kaki putih antara lain ketahanan terhadap fluktuasi kualitas air, pertumbuhan yang relatif cepat, dan kemampuan ditebar pada kepadatan tinggi, sehingga memberikan manfaat bagi petambak. Kabupaten Lamongan merupakan salah satu daerah potensial untuk mengembangkan budidaya udang kaki putih karena luas tambaknya yang luas dan jumlah petambak yang banyak. Sentra budidaya terbesar berada di Kecamatan Karangbinayun, tepatnya di Desa Pendowolimo yang terkenal dengan budidaya udang kaki putih yang aktif.

Namun, berbagai isu masih sering dibahas oleh petani, seperti ketidakstabilan iklim, permintaan konsumen, dan suhu air [3]. Kondisi kolam, termasuk luas lahan, jumlah pakan, dan penanaman benih, juga berdampak signifikan pada hasil panen. Di Desa Pendowolimo, Kecamatan Karangbinayun, proses perencanaan dan pendataan budidaya udang kaki putih masih dilakukan secara manual, sehingga mengakibatkan pengendalian panen yang buruk bahkan berisiko gagal panen. Pengendalian dan pengawasan panen sangat penting untuk memaksimalkan produksi dan meningkatkan keuntungan. Penelitian sebelumnya tentang pengembangan budidaya udang kaki putih di Desa Sidokumpul, Kecamatan Lamongan, menekankan perlunya strategi untuk meningkatkan hasil panen melalui optimalisasi budidaya, seperti peningkatan unit pembenihan, peningkatan sarana dan prasarana, serta peningkatan kualitas teknis, moral, dan spiritual sumber daya petani [4].

Algoritma KNN unggul karena kuat terhadap data pelatihan yang berisik dan mampu menangani kumpulan data yang besar. Namun, kelemahannya termasuk menentukan nilai K, memilih jenis jarak dan atribut yang memengaruhi akurasi, dan biaya komputasi yang tinggi karena perhitungan jarak dilakukan pada seluruh data pelatihan [5].

Beberapa penelitian sebelumnya relevan dengan masalah ini. Penelitian [6] berjudul "Penerapan Metode KNN dalam Memprediksi Hasil Panen Tebu di Kabupaten Takalar" menunjukkan bahwa data baru diklasifikasikan berdasarkan mayoritas kategori KNN, dengan akurasi tertinggi 76,92% pada nilai K=7 dari 13 data pelatihan. Penelitian [7] berjudul "Penerapan metode k-nearest neighbor untuk memprediksi penjualan sepeda motor terlaris" menggunakan data penjualan PT. Sumber Rejeki Jabar periode Januari–Desember 2022, dengan akurasi tertinggi 96,15% pada nilai K 5. Riset [8] berjudul "Pendekatan Regresi KNN untuk Prediksi Saham GameStop", menunjukkan parameter

terbaik pada  $K=3$  dengan MSE 0,1 dan  $R^2$  0,98, menunjukkan akurasi tinggi. Penelitian ini menunjukkan efektivitas KNN Regression untuk analisis prediktif saham GameStop dan memiliki potensi pengembangan lebih lanjut.

Mengacu pada pernyataan di atas, penulis memotivasi penelitian di daerah tersebut, berjudul "Sistem Prediksi Panen Udang Kaki Putih Menggunakan Metode K-Nearest Neighbor." Sistem prediksi K-Nearest Neighbor ini memprediksi hasil panen udang kaki putih di masa depan dengan memanfaatkan data historis. Sistem prediksi ini digunakan untuk menentukan strategi pengelolaan di masa depan untuk memaksimalkan hasil panen.

## 2. Metodologi

### 2.1. Metode Analisis

#### 2.1.1. K-Metode Tetangga Terdekat

K-Nearest Neighbors (KNN) adalah teknik pembelajaran yang diawasi yang menggunakan data pelatihan untuk mengklasifikasikan data baru berdasarkan kedekatannya dengan data pelatihan yang ada. Algoritma KNN mengklasifikasikan data baru berdasarkan kedekatannya dengan data pelatihan, menggunakan berbagai metrik jarak seperti jarak Euclidean, jarak Manhattan, dan jarak Minkowski[9]. Menurut [10] KNN adalah algoritma yang bekerja berdasarkan kategori mayoritas tetangga terdekat, yang didasarkan pada jarak minimum.

KNN mengklasifikasikan data baru berdasarkan tetangga terdekat dengan menghitung jarak terpendek antara data kueri dan pelatihan, di mana penelitian ini menggunakan jarak Euclidean.

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

Deskripsi:

$x_i$  = Data dalam dimensi ke- $i$

$y_i$  = Data referensi dalam dimensi ke- $i$

$n$  = Jumlah dimensi atau fitur data

#### 2.1.2. Rata-rata Kesalahan Persentase Absolut (MAPE)

Mean Absolute Percentage Error (MAPE) adalah ukuran statistik yang digunakan untuk mengevaluasi keakuratan hasil perkiraan dengan menghitung persentase rata-rata kesalahan absolut. Nilai MAPE menunjukkan tingkat kesalahan prediksi dibandingkan dengan nilai aktual, di mana MAPE yang lebih kecil mewakili akurasi perkiraan yang lebih tinggi[11].

$$\frac{1}{n} \sum \frac{|D_t - y_t|}{D_t} \times 100\% \quad (2)$$

Deskripsi:

$D_t$  = Nilai aktual

$Y_t$  = Nilai yang diprediksi

$n$  = Jumlah data

Nilai MAPE dapat diartikan ke dalam kategori, yaitu [12] :

<10% = sangat akurat

10-20% = cukup akurat

20-50% = wajar

>50% = tidak akurat

## 2.2. Tahapan Penelitian

### 2.2.1. Teknik Pengumpulan Data

Menurut [13], teknik pengumpulan data dalam penelitian ini dilakukan melalui wawancara dengan balai desa dan masyarakat pemilik tambak, termasuk Kepala Desa Pendowolimo dan Bapak M. Iskandar sebagai sumber utama, pengamatan langsung ke lokasi tambak untuk mencatat kondisi dan pengelolaan, serta analisis dokumen berupa laporan panen yang berisi informasi masa panen, luas tambak, jumlah bibit, masa pemeliharaan, pakan, dan data lain yang terkait dengan budidaya udang Vaname.

### 2.2.2. Analisis Data

Data yang digunakan untuk pengujian sistem adalah data penjualan historis dari studi kasus terkait jumlah 150 data panen, dalam format datatambakubed.xlsx di Microsoft Excel, yang diperoleh pada 9 Januari 2025. Setelah pengamatan dan wawancara, data kemudian disajikan dalam tabel berikut.

Tabel 1. Data Wawancara

No	Tahun	Waktu Panen	Luas Kolam (m <sup>2</sup> )	Tanaman (rean)	Durasi Pemeliharaan (Hari)	Memberi makan (kg)	Hasil (Kg)
1	2020	panen 1	500	20	60	70	95
2	2020	panen 2	500	15	50	50	80
3	2020	panen 3	500	5	55	40	52
4	2020	panen 1	600	20	50	80	125
5	2020	panen 2	600	10	52	50	90

6	2020	panen 3	600	7	56	50	75
7	2020	panen 1	550	17	56	65	84
No	Tahun	Waktu Panen	Luas Kolam (m <sup>2</sup> )	Tanaman (rean)	Durasi Pemeliharaan (Hari)	Memberi makan (kg)	Hasil (Kg)
8	2020	panen 2	550	12	55	50	73
9	2020	panen 3	550	7	60	35	50
10	2020	panen 1	225	12	53	55	65
...	...	...	...	...	...	...	...
149	2024	panen 2	150	2	48	5	10
150	2024	panen 3	150	2	48	6	8

### 2.2.3. Langkah Kerja

- Analisis dan pengumpulan data  
Untuk mengontrol dan mengelola hasil panen udang vaname yang terkadang tidak memenuhi harapan, data tambak udang dikumpulkan dengan beberapa variabel pendukung, data diperoleh dengan beberapa variabel pendukung seperti yang ditunjukkan pada tabel berikut.

Tabel 2. Data Mentah

No	Tahun	Waktu Panen	Luas Kolam (m <sup>2</sup> )	Tanaman (rean)	Durasi Pemeliharaan (Hari)	Memberi makan (kg)	Hasil (Kg)
1	2020	panen 1	500	20	60	70	95
2	2020	panen 2	500	15	50	50	80
3	2020	panen 3	500	5	55	40	52
4	2020	panen 1	600	20	50	80	125
5	2020	panen 2	600	10	52	50	90
6	2020	panen 3	600	7	56	50	75
7	2020	panen 1	550	17	56	65	84
8	2020	panen 2	550	12	55	50	73
9	2020	panen 3	550	7	60	35	50
10	2020	panen 1	225	12	53	55	65

- Pra-pemrosesan data  
Tahap pengolahan data dilakukan untuk membersihkan, menormalkan, dan mengubah format agar siap digunakan dalam analisis. Proses ini meliputi penghapusan nilai kosong, penghapusan kolom yang tidak perlu seperti angka dan tahun, serta perubahan variabel menjadi X1 (masa panen), X2 (luas tambak), X3 (penanaman), X4 (masa pemeliharaan), dan X5 (pakan). Data nonnumerik kemudian diubah menjadi angka, misalnya Panen 1 = 30, Panen 2 = 20, Panen 3 = 10, dan kemudian normalisasi dilakukan dengan menggunakan metode Min-Max sehingga semua fitur berada pada skala yang sama. Berikut adalah rumus min-max :

$$\hat{x} = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{3}$$

Deskripsi:

X = data

Min(x) = data terkecil

Maks(x) = data terbesar

Tabel 3. Dataset

	X1	X2	X3	X4	X5	Y
1	0,733333333	1	1	0,868421	95	
0,5	0,733333333	0,73684211	0,333333333	0,605263	80	
0	0,733333333	0,21052632	0,666666667	0,473684	52	
1	1	1	0,333333333	1	125	
0,5	1	0,47368421	0,466666667	0,605263	90	
0	1	0,31578947	0,733333333	0,605263	75	
1	0,866666667	0,84210526	0,733333333	0,802632	84	
0,5	0,866666667	0,57894737	0,666666667	0,605263	73	
0	0,866666667	0,31578947	1	0,407895	50	
1	0	0,57894737	0,533333333	0,671053	65	

Tabel di atas menampilkan hasil data pra-pemrosesan dan normalisasi, dimana data yang disusun adalah data yang dibutuhkan dan relevan untuk perhitungan, yang disebut dengan dataset.

- Perhitungan Jarak  
Siapkan input/data yang diperlukan untuk perhitungan. Misalnya, masukan:  
X1 (Periode Panen) = Panen 1  
X2 (Luas Kolam) = 600m<sup>2</sup>  
X3 (Penanaman) = 3 panen  
X4 (Periode Pemeliharaan) = 55 hari

X5 (Pakan) = 90 kg

Jarak Euclidean dihitung menggunakan persamaan 1.

1. Pelaksanaan

a. Jarak ke baris pertama

$$d_1 = \sqrt{(1-1)^2 + (0,818-1)^2 + (1-0)^2 + (1-0,5)^2 + (0,75-0,868)^2}$$

$$d_1 = \sqrt{0 + 0,0331 + 1 + 0,25 + 0,01392} = \sqrt{1,29702} = 1,13886$$

b. Jarak ke baris kedua

$$d_2 = \sqrt{(0,5-1)^2 + (0,818-1)^2 + (0,737-0)^2 + (0,333-0,5)^2 + (0,605-0,875)^2}$$

$$d_2 = \sqrt{0,25 + 0,331 + 0,0543 + 0,0278 + 0,0729} = \sqrt{0,92702} = 0,9629$$

c. Jarak ke baris ketiga

$$d_3 = \sqrt{(0-1)^2 + (0,818-1)^2 + (0,211-0)^2 + (0,667-0,5)^2 + (0,474-0,875)^2}$$

$$d_3 = \sqrt{1 + 0,0331 + 0,04452 + 0,0278 + 0,1608} = \sqrt{1,26633} = 1,1253$$

d. Jarak ke baris keempat

$$d_4 = \sqrt{(1-1)^2 + (1-1)^2 + (1-0)^2 + (0,333-0,5)^2 + (1-0,875)^2}$$

$$d_4 = \sqrt{0 + 0 + 1 + 0,0278 + 0,01562} = \sqrt{1,043514} = 1,0215$$

e. Jarak ke baris kelima

$$d_5 = \sqrt{(0,5-1)^2 + (1-1)^2 + (0,474-0)^2 + (0,467-0,5)^2 + (0,605-0,875)^2}$$

$$d_5 = \sqrt{0,25 + 0 + 0,2246 + 0,00108 + 0,0729} = \sqrt{0,548665} = 0,7407$$

f. Jarak ke baris keenam

$$d_6 = \sqrt{(0-1)^2 + (1-1)^2 + (0,316-0)^2 + (0,733-0,5)^2 + (0,605-0,875)^2}$$

$$d_6 = \sqrt{1 + 0 + 0,09985 + 0,054289 + 0,0729} = \sqrt{1,22704} = 1,1077$$

g. Jarak ke baris ketujuh

$$d_7 = \sqrt{(1-1)^2 + (0,909-1)^2 + (0,482-0)^2 + (0,733-0,5)^2 + (0,75-0,875)^2}$$

$$d_7 = \sqrt{0 + 0,00828 + 0,2323 + 0,05428 + 0,01562} = \sqrt{0,310517} = 0,5572$$

h. Jarak ke baris kedelapan

$$d_8 = \sqrt{(0,5-1)^2 + (0,909-1)^2 + (0,579-0)^2 + (0,667-0,5)^2 + (0,803-0,875)^2}$$

$$d_8 = \sqrt{0,25 + 0,00828 + 0,3352 + 0,02788 + 0,00518} = \sqrt{0,62659} = 0,7916$$

i. Jarak ke baris kesembilan

$$d_9 = \sqrt{(0-1)^2 + (0,909-1)^2 + (0,316-0)^2 + (1-0,5)^2 + (0,605-0,875)^2}$$

$$d_9 = \sqrt{1 + 0,008281 + 0,09985 + 0,25 + 0,0729} = \sqrt{1,431037} = 1,1963$$

j. Jarak ke baris kesepuluh

$$d_{10} = \sqrt{(1-1)^2 + (0,318-1)^2 + (0,579-0)^2 + (0,533-0,5)^2 + (0,408-0,875)^2}$$

$$d_{10} = \sqrt{0 + 0,46512 + 0,33524 + 0,00108 + 0,21808} = \sqrt{1,019543} = 1,0097$$

Dari semua perhitungan diperoleh hasil sebagai berikut :

Tabel 4. Hasil K

Garis	Nilai Euclidean	Hasil	K
1	1,1389	125	4
2	0,9629	80	5
3	1,1253	45	9
4	1,0215	135	7
5	0,7407	90	3
6	1,1077	65	6
7	0,5572	110	2
8	0,7916	80	1
9	1,1963	50	10
10	1,0097	90	8

Berdasarkan tabel di atas, nilai K-3 terdekat diperoleh di baris pertama, baris ketujuh dan baris keempat. Berdasarkan masukan masa panen pertama, dengan luas kolam 600m<sup>2</sup> dan penanaman 3 bibit, dengan masa pemeliharaan 55 hari dan pemberian pakan 90kg, diprediksi dengan mengambil nilai y atau hasil panen dari kolom pertama, ketujuh, dan keempat sebagai berikut : Berikut ini adalah persamaan untuk menghitung rata-rata [14]:

$$Y = \frac{1}{k} \sum_{i=1}^k Y_i \quad (4)$$

Implementasi nilai input dan hasil perhitungan

$$\text{Prediksi} = \frac{80 + 110 + 90}{3}$$

$$\text{Prediksi} = \frac{280}{3}$$

Prediksi = 93,3

Jadi hasil yang diprediksi adalah 93,3 kg

**2.2.4. MAPE Kalkulasi**

Perhitungan evaluasi menggunakan Mean Absolute Percentage Error (MAPE) dilakukan untuk menentukan tingkat kesalahan perhitungan berdasarkan data penjualan yang ada. Perhitungan evaluasi MAPE memerlukan renormalisasi kumpulan data sebelumnya.

Tabel 5. Kalkulasi MAPE

Garis	Hasil Panen (y)=(Dt)	Hasil Prediksi (Ft)	(Dt-Ft)/Dt
1	125	95	93.00
2	80	80	77.00
Garis	Hasil Panen (y)=(Dt)	Hasil Prediksi (Ft)	(Dt-Ft)/Dt
3	45	52	53.67
4	135	125	115.00
5	90	90	89.00
6	65	75	73.67
7	110	84	79.67
8	80	73	75.00
9	50	50	52.33
10	90	65	63.00
Σ	789	771,3	0,35591

Kolom hasil variabel (y) diubah menjadi Dt, kolom Hasil Prediksi disebut variabel (Ft), dan perhitungan nilai kesalahan untuk setiap data adalah sebagai berikut.

$$\frac{(Dt - Ft)}{Dt} \tag{5}$$

Nilai kesalahan dijumlahkan untuk menghitung MAPE keseluruhan.

Setelah memperoleh hasil perhitungan kesalahan pada Tabel 3.5, MAPE dihitung menggunakan rumus 2.

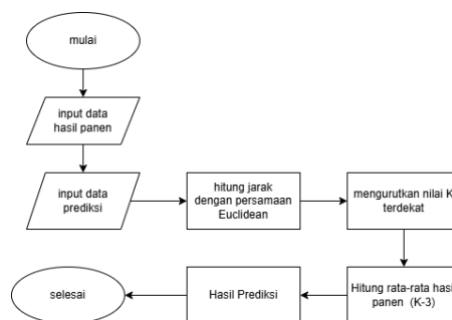
$$MAPE = \frac{0,35591}{10} \times 100\% = 3,5\%$$

Nilai kesalahan MAPE adalah 3,5%. Angka ini dinilai sangat akurat dalam perhitungan MAPE.

**2.2.5. Desain Proses**

Desain proses ini merupakan desain sistem untuk penelitian ini, yang menggunakan pendekatan berbasis proses. Oleh karena itu, dalam desain sistem, peneliti menerapkan DFD (Data Flow Diagram). DFD digunakan sebagai alat untuk memodelkan aliran data dan proses dalam sistem. Proses yang digunakan dalam desain sistem ini termasuk membuat diagram alur dan diagram level 0 dan 1.

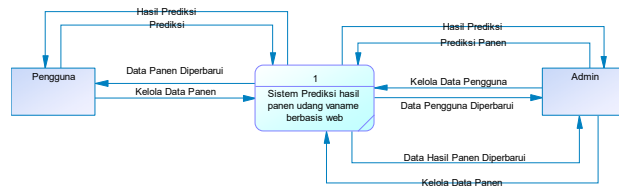
1. Diagram Alur K-Tetangga Terdekat (KNN)  
Berikut ini adalah langkah-langkah proses dalam algoritma K-Nearest Neighbors (KNN), khususnya dalam penerapannya untuk memprediksi data penjualan. Langkah-langkah ini meliputi input data, perhitungan jarak, pemilihan tetangga terdekat, dan hasil prediksi dan evaluasi [15]. seperti yang ditunjukkan pada gambar di bawah ini.



Gambar 1: Diagram Alur Prediksi Udang Vaname Menggunakan Metode KNN

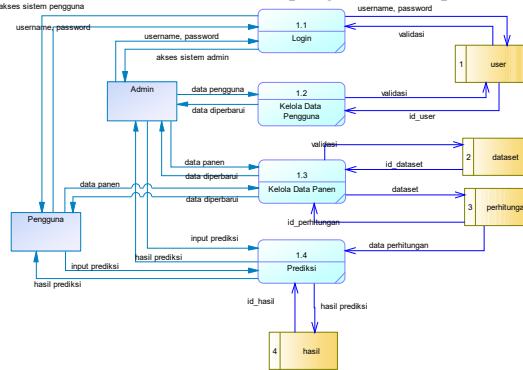
Diagram alur pada gambar di atas menggambarkan tahapan algoritma K-Nearest Neighbors (KNN), dari awal hingga akhir. Prosesnya dimulai dengan memasukkan data yang diperlukan, yaitu data pelatihan dan pengujian. Selanjutnya, data pelatihan disiapkan dalam array untuk menyederhanakan pemrosesan. Setelah itu, jarak antara data pengujian dan setiap kumpulan data pelatihan dihitung menggunakan persamaan matematika, seperti Jarak Euclidean. Jarak yang dihitung kemudian diurutkan dari terkecil ke terbesar untuk menentukan tetangga terdekat. Berdasarkan jumlah tetangga yang telah ditentukan (K), sistem memilih data terdekat dan menghitung rata-rata nilainya untuk menghasilkan prediksi. Setelah hasil prediksi diperoleh, langkah selanjutnya adalah menghitung akurasi dengan membandingkan hasil prediksi dengan data aktual. Proses ini berakhir setelah semua tahapan selesai. Diagram alur ini memberikan gambaran yang jelas tentang cara kerja algoritma KNN.

1. Diagram Alur Data (DFD)  
Data Flow Diagram adalah jaringan yang menggambarkan sistem otomatis atau terkomputerisasi, manualisasi, atau kombinasi keduanya, yang deskripsinya disusun dalam bentuk kumpulan komponen sistem yang saling berhubungan sesuai dengan aturan main.



Gambar 2: Diagram Alur Data Level 0

Diagram Aliran Data Level 0 untuk sistem prediksi panen udang kaki putih berbasis web untuk tambak di Desa Pendowolimo menggambarkan aliran data secara keseluruhan. Sistem ini berinteraksi dengan dua entitas eksternal: karyawan dan administrator. Karyawan mengirimkan data penjualan, termasuk informasi seperti nama produk dan jangka waktu, ke sistem. Data yang diterima oleh sistem kemudian diolah menggunakan algoritma K-Nearest Neighbors (K-NN) untuk menghasilkan prediksi panen. Hasil prediksi ini dapat diakses oleh administrator untuk membantu pengambilan keputusan strategis dalam budidaya udang kaki putih.



Gambar 3: Diagram Alur Data Level 1

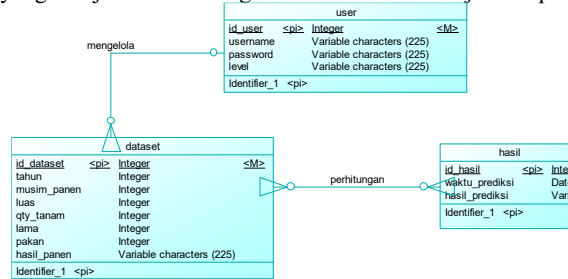
DFD Level 1 terdiri dari empat proses utama, yaitu Login (1.1) untuk otentikasi pengguna, Kelola Data Pengguna (1.2) bagi admin untuk menambahkan, memperbarui, atau menghapus data pengguna, Mengelola Data Penjualan (1.3) untuk mencatat dan mengelola transaksi dengan validasi dari entitas terkait, dan Prediksi (1.4) yang memanfaatkan data penjualan historis untuk menghasilkan perkiraan permintaan. Admin memiliki kendali penuh atas seluruh proses, sedangkan karyawan hanya memiliki hak akses terbatas sesuai dengan peran mereka.

2.2.6. Disain Database

Desain data ini bertujuan untuk menentukan konsep desain atau rencana yang akan diterapkan pada sistem yang sedang dibangun. Proses ini menjelaskan pembuatan ikhtisar database yang mendukung operasi sistem, dengan tujuan merancang sistem untuk memprediksi angka panen vaname menggunakan metode K-Nearest Neighbors (K-NN). Berikut ini adalah database yang digunakan sebagai acuan sistem.

1. CDM (Model Data Konseptual)

CDM adalah konsep yang terkait dengan perspektif pengguna terhadap data yang tersimpan dalam database. CDM dibuat dalam bentuk tabel tanpa tipe data yang menjelaskan hubungan antar tabel untuk tujuan implementasi dalam database.

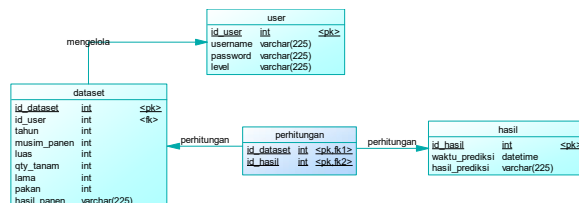


Gambar 4: Model Data Konseptual

Gambar di atas menggambarkan hubungan database yang digunakan dalam sistem prediksi panen vaname menggunakan metode K-Nearest Neighbors (K-NN). Ada tiga tabel, masing-masing dengan perannya sendiri: tabel pengguna untuk aktivitas login, tabel himpunan data untuk data panen dan kebutuhan perhitungan, dan tabel hasil untuk menyimpan data prediksi.

1. PDM (Model Data Fisik)

PDM adalah konsep yang menjelaskan detail bagaimana data disimpan dalam database. PDM adalah representasi fisik dari desain database yang siap diimplementasikan dalam DBMS, sehingga nama tabel juga merupakan nama sebenarnya dari tabel yang siap diimplementasikan.



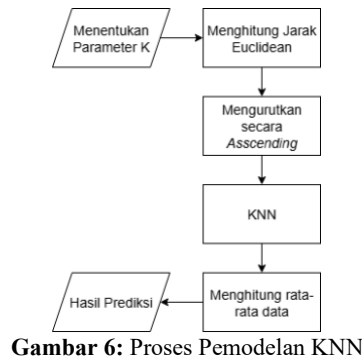
Gambar 5: Model Data Fisik

Gambar di atas menjelaskan hubungan database yang digunakan dalam sistem prediksi panen vaname menggunakan metode K-Nearest Neighbors (K-NN).

### 3. Hasil dan Pembahasan

#### 3.1. KNN Modeling Process

Gambar berikut menunjukkan tahapan proses pemodelan K-Nearest Neighbor (KNN) menggunakan Excel.



Gambar 6: Proses Pemodelan KNN

#### 3.2. Pemodelan Penelitian

##### 3.2.1. Menentukan Nilai K

Menentukan nilai tetangga terdekat K. Untuk mendapatkan hasil yang lebih bervariasi, beberapa perhitungan tetangga dilakukan menggunakan K=3; K=5, atau K=7.

##### 3.2.2. Menghitung Jarak Euclidean

Perhitungan Jarak Euclidean untuk setiap objek terhadap data sampel yang diberikan. Percobaan ini menggunakan variabel pengujian atau input sebagai berikut:

- X1 (Periode Panen) = Panen 1
- X2 (Luas Kolam) = 500m<sup>2</sup>
- X3 (Penanaman) = 6 panen
- X4 (Periode Pemeliharaan) = 65 hari
- X5 (Pakan) = 80 kg

Pada tabel di bawah ini terdapat perhitungan menggunakan rumus persamaan Jarak Euclidean.

Tabel 6. Hitung Jarak Euclidean

<i>N</i> <i>o</i>	<i>X</i> <i>1</i>	<i>X</i> <i>2</i>	<i>X</i> <i>3</i>	<i>X</i> <i>4</i>	<i>X</i> <i>5</i>	<i>Kalkulasi</i>	<i>Euclidean</i>
1	1	0.818	1	1	0.868	$\sqrt{(1-1)^2 + (0,818-0,818)^2 + (1-0,263)^2 + (1-1)^2 + (0,75-0,938)^2}$	608.871
2	0.5	0.818	0.736	0.333	0.605	$\sqrt{(0,5-1)^2 + (0,818-0,818)^2 + (0,737-0,263)^2 + (0,333-1)^2 + (0,605-0,938)^2}$	608.994
3	0	0.818	0.210	0.666	0.473	$\sqrt{(0-1)^2 + (0,818-0,818)^2 + (0,211-0,263)^2 + (0,667-1)^2 + (0,474-0,938)^2}$	609.010
4	1	1	1	0.333	1	$\sqrt{(1-1)^2 + (1-0,818)^2 + (1-6)^2 + (0,333-1)^2 + (0,875-0,938)^2}$	608.733
5	0.5	1	0.473	0.466	0.605	$\sqrt{(0,5-1)^2 + (1-0,818)^2 + (0,474-0,263)^2 + (0,467-1)^2 + (0,605-0,938)^2}$	608.805
...	...	...	...	...	...	...	...
1460.5	0.090	0.105	0.2	0.184		$\sqrt{(0,5-1)^2 + (0,090-0,818)^2 + (0,105-0,263)^2 + (0,2-1)^2 + (0,184-0,938)^2}$	609.786
1470	0.090	0.052	0.2	0		$\sqrt{(0-1)^2 + (0,090-0,818)^2 + (0,052-0,263)^2 + (0,2-1)^2 + (0-0,938)^2}$	609.838

1481	0.181	0.105	0.2	0.105	$\sqrt{(1-1)^2 + (0,181-0,818)^2 + (0,105-0,263)^2 + (0,2-1)^2 + (0,105-0,938)^2}$	609.685
1490.5	0.181	0.052	0.2	0.013	$\sqrt{(0,5-1)^2 + (0,181-0,818)^2 + (0,052-0,263)^2 + (0,2-1)^2 + (0,013-80)^2}$	609.722
1500	0.181	0.052	0.2	0.026	$\sqrt{(0-1)^2 + (0,181-0,818)^2 + (0,052-0,263)^2 + (0,2-1)^2 + (0,026-0,938)^2}$	609.745

Pada tabel di atas, perhitungan jarak Euclidean menggunakan rumus 150 data pelatihan pada dataset yang diuji dengan data input (testing).

### 3.2.3. Mengurutkan hasil perhitungan dalam urutan menaik

Proses selanjutnya adalah mengurutkan data yang dihitung untuk memudahkan pengguna menemukan tetangga terdekat pada tabel di bawah ini.

Tabel 7. Klasifikasi KNN

Euclidean	Jarak	KNN	Klasifikasi
608.871	1	Ya	1
608.994	2	Ya	2
608.733	3	Ya	3
608.733	4	Tidak	-
608.742	5	Tidak	-
...	...	...	...
609.888	146	Tidak	-
609.898	147	Tidak	-
609.912	148	Tidak	-
609.926	149	Tidak	-
609.926	150	Tidak	-

### 3.2.4. Average Classification Results

Berdasarkan hasil klasifikasi KNN menggunakan K=3 diperoleh 3 data tetangga terdekat, yang ditampilkan pada tabel di bawah ini.

Table 8. Classification Results Data

No	X1	X2	X3	X4	X5	Harvest Amount	Classification
34	1	1.000	0.684	1.000	0.868	130	1
64	1	1.000	0.842	0.667	0.934	130	2
94	1	1.000	0.737	0.667	0.803	98	3

Untuk menghitung rata-rata menggunakan rumus rata-rata dan menghasilkan prediksi berikut

$$\text{Prediction} = \frac{130 + 130 + 98}{3}$$

Harvest Prediction = 119,3 Riwayat medis

### 3.3. Implementasi Sistem

Implementasi sistem adalah tahap implementasi dari hasil desain dari bab sebelumnya. Pada tahap ini, desain sistem, termasuk desain basis data, alur sistem, dan desain antarmuka pengguna, diwujudkan dalam bentuk program yang dapat dieksekusi. Implementasi dilakukan dengan menggunakan bahasa pemrograman dan perangkat lunak pendukung yang sesuai dengan kebutuhan sistem.

Hasil implementasi ditampilkan melalui tampilan antarmuka (tangkapan layar) dari setiap halaman sistem, bersama dengan cuplikan kode utama yang mewakili fungsi penting sistem. Dengan implementasi ini, sistem prediksi panen udang kaki putih berbasis metode K-Nearest Neighbors (K-NN) dapat beroperasi sesuai rencana.

### 3.3.1. Implementasi Halaman Himpunan Data

No	Tahun Panen	Musim Panen	Luas Tambak	Nanam (Rean)	Lama Pemeliharaan	Pakan	Hasil Panen	Aksi
1	2020	Musim 3	500 M <sup>2</sup>	15	60 Hari	80 Kg	150 Kg	[Edit] [Hapus]
2	2020	Musim 2	500 M <sup>2</sup>	10	50 Hari	50 Kg	80 Kg	[Edit] [Hapus]
3	2020	Musim 1	500 M <sup>2</sup>	7	55 Hari	20 Kg	50 Kg	[Edit] [Hapus]
4	2020	Musim 3	600 M <sup>2</sup>	20	50 Hari	100 Kg	200 Kg	[Edit] [Hapus]
5	2020	Musim 2	600 M <sup>2</sup>	10	52 Hari	50 Kg	90 Kg	[Edit] [Hapus]
6	2020	Musim 1	600 M <sup>2</sup>	7	56 Hari	50 Kg	65 Kg	[Edit] [Hapus]
7	2020	Musim 3	550 M <sup>2</sup>	15	51 Hari	80 Kg	120 Kg	[Edit] [Hapus]
8	2020	Musim 2	550 M <sup>2</sup>	10	55 Hari	50 Kg	80 Kg	[Edit] [Hapus]
9	2020	Musim 1	550 M <sup>2</sup>	7	60 Hari	20 Kg	50 Kg	[Edit] [Hapus]
10	2020	Musim 3	225 M <sup>2</sup>	15	53 Hari	50 Kg	90 Kg	[Edit] [Hapus]

Gambar 7: Himpunan Data Halaman

Halaman dataset pada gambar di atas menampilkan semua data historis panen udang kaki putih. Data ini meliputi tahun, masa panen, luas kolam, jumlah penanaman, periode pemeliharaan, jumlah pakan, dan hasil panen. Halaman ini berfungsi sebagai pusat manajemen data utama sebelum membuat prediksi.

### 3.3.2. Implementasi Halaman Prediksi

Gambar 8: Halaman Prediksi

Gambar di atas menampilkan halaman yang menyediakan formulir untuk input data prediktif, termasuk luas tambak, volume tanam, durasi pemeliharaan, jumlah pakan, tahun, dan waktu panen. Setelah pengguna memasukkan data dan menekan tombol prediksi, sistem memproses data menggunakan metode KNN dan menampilkan hasilnya.

### 3.3.3. Implementasi Halaman Hasil Prediksi

Hasil Perhitungan Prediksi KNN (dengan \*\*normalisasi Min-Max\*\*) dari inputan berikut:

Input Asli:

- Luas Tambak (A<sup>2</sup>) = 500
- Tanam (Rean) = 8
- Lama Panen (Hari) = 60
- Banyak Pakan (Kg) = 100
- Musim Panen = 10

Setelah Normalisasi (0-1):

- Luas = 0.8182
- Tanam = 0.3684
- Lama = 1
- Pakan = 1
- Musim Panen = 0

Nilai K = 3  
Hasil Prediksi: 53 Kg

---

1. Semua Perhitungan Jarak (dengan data normalisasi)

Perhitungan jarak menggunakan rumus Euclidean pada data yang telah dinormalisasi:  $d(x,y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$

No	LUAS		QTY TANAM		LAMPA		PAKAN		MUSIM PANEN		Tahun	Total Jarak (J)	Hasil Panen
	Norm.	(x-y) <sup>2</sup>	Norm.	(x-y) <sup>2</sup>	Norm.	(x-y) <sup>2</sup>	Norm.	(x-y) <sup>2</sup>	Norm.	(x-y) <sup>2</sup>			
1	1	0.03058	0.3158	0.00277	0.7333	0.071111	0.5679	0.186709	0	0	2020	0.541893	65
2	0.9091	0.008264	0.3158	0.00277	1	0	0.3827	0.381039	0	0	2020	0.626158	50

2. 3 Tetangga Terdekat (berdasarkan normalisasi)

No	LUAS (Norm.)	QTY TANAM (Norm.)	LAMA (Norm.)	PAKAN (Norm.)	MUSIM PANEN (Norm.)	Tahun	Jarak	Hasil Panen
1	1	0.3158	0.7333	0.5679	0	2020	0.541893	65
2	0.9091	0.3158	1	0.3827	0	2020	0.626158	50
3	0.8182	0.2105	0.6667	0.4444	0	2020	0.666846	45

3. Hasil Prediksi

Berdasarkan metode K-Nearest Neighbor (menggunakan data normalisasi Min-Max) dengan K = 3, maka hasil prediksi panen adalah:

$$Y_{\text{prediksi}} = \frac{1}{3} (65 + 50 + 45) = 53.3333 \text{ Kg}$$

Gambar 9: Halaman Hasil Prediksi

Gambar di atas menunjukkan hasil akhir dari prediksi panen udang kaki putih dalam kilogram. Sistem ini juga menampilkan nilai tetangga terdekat dan perhitungan matematisnya.

### 3.3.4. Implementasi Halaman Evaluasi

Evaluasi Prediksi Menggunakan Metode MAPE								
Luas Tambak	Jumlah Naman	Lama Pemeliharaan	Banyak Pakan	Tahun	Musim Panen	Hasil Panen	Hasil Prediksi	Kesalahan
500	20	60	80	2020	30	125	121,6667	0,0267
500	15	50	50	2020	20	80	83	0,0375
500	5	55	40	2020	10	45	50	0,1111
600	20	50	80	2020	30	125	125	0,0741
600	10	52	50	2020	20	90	78	0,1333
600	7	56	50	2020	10	65	53,3333	0,1795
550	15	60	70	2020	30	110	108,3333	0,0552
550	10	55	50	2020	20	80	73,3333	0,0833
550	7	60	35	2020	10	50	53,3333	0,0667
225	15	53	50	2020	30	90	80	0,1111
						Total MAPE (Mean Absolute Percentage Error):		19,48 %

Gambar 10: Halaman Evaluasi

Halaman ini menyajikan hasil perhitungan evaluasi menggunakan metode MAPE. Gambar di atas menampilkan perbandingan antara nilai aktual dan prediksi untuk sekumpulan kumpulan data, lengkap dengan persentase kesalahan. Informasi ini digunakan untuk menilai kinerja keseluruhan sistem prediksi. Ada juga fitur untuk mengubah nilai K, yang digunakan untuk menilai keakuratan data menggunakan tetangga K lainnya sebagai uji akurasi.

## 4. Kesimpulan

Berdasarkan penelitian dan implementasi sistem prediksi panen udang vaname menggunakan metode K-Nearest Neighbors (KNN), sistem tersebut berhasil dikembangkan dengan memanfaatkan data historis (periode panen, luas tambak, jumlah benih, masa pemeliharaan, dan jumlah pakan) dan perhitungan jarak Euclidean dengan nilai  $K = 3$  yang dinilai optimal. Sistem ini dilengkapi dengan fitur CRUD dan visualisasi hasil prediksi pada antarmuka web. Dari pengujian menggunakan Mean Absolute Percentage Error (MAPE), diperoleh nilai 9,6% yang termasuk dalam kategori sangat baik, sehingga sistem dinilai akurat dan mampu memberikan hasil yang mendekati kenyataan bahkan dengan data historis yang terbatas namun terstruktur.

## Pengakuan

Dengan rasa syukur, penulis mengungkapkan rasa syukurnya kepada Allah SWT atas rahmat dan keberkahan-Nya, memungkinkan keberhasilan penyelesaian penelitian skripsi ini, bertajuk "Sistem Prediksi Panen Udang Vaname Menggunakan Metode K-Tetangga Terdekat (KNN)." Penulis juga mengucapkan terima kasih kepada orang tua dan pembimbingnya atas bimbingan dan arahan mereka selama penyusunan skripsi ini. Lebih lanjut, ia mengucapkan terima kasih kepada semua pihak terkait atas dukungan dan kerjasamanya, terutama yang telah memberikan data dan kesempatan untuk melakukan penelitian ini.

## Referensi

- [1] A. Jayadi, D. Susiati, and I. K. Tjahjani, "Peramalan Penjualan Ikan Dengan Simulasi Double Exponential Smoothing Di Lamongan ( Studi Kasus Tambak Ikan Bapak Wiharnoto )," *Pros. Semin. Nas. Tek. Ind. [Snti]*, vol. 5, pp. 1–9, 2023.
- [2] M. . Sukadi, "Vannamei, Fenomena Baru Dalam Bisnis Budidaya Udang," *Bul. Dep. Kelaut. dan Perikanan.*, 2004.
- [3] Muzahar, *Teknologi dan Manajemen Budidaya Udang*. Tanjungpinang: UMRAH PRESS, 2020.
- [4] K. Laila, U. Siregar, A. B. Sinaga, D. A. A. Marpaung, R. Handayani, and B. Wahyudi, "Teknik Pembesaran Udang Vaname ( Litopenaeus vannamei ) di Desa Pesisir , Kecamatan Talawi , Kabupaten Batu Bara Vaname Shrimp Enlargement Technique ( Litopenaeus vannamei ) in Pesisir Village , Talawi District , Batu Bara Regency," *Grouper*, vol. 14, no. 2, pp. 112–117, 2023.
- [5] S. R. Cholil, T. Handayani, R. Prathivi, and T. Ardianita, "Implementasi Algoritma Klasifikasi K-Nearest Neighbor (KNN) Untuk Klasifikasi Seleksi Penerima Beasiswa," *IJCIT (Indonesian J. Comput. Inf. Technol.)*, vol. 6, no. 2, pp. 118–127, 2021, doi: 10.31294/ijcit.v6i2.10438.
- [6] N. Situju, L. N. Hayati, and W. Astuti, "Penerapan Metode KNN dalam Memprediksi Hasil Panen Kebun Tebu di Kab Takalar," *Bul. Sist. Inf. dan Teknol. Islam*, vol. 4, no. 1, pp. 35–40, 2023, doi: 10.33096/busiti.v4i1.1474.
- [7] R. Rismala, I. Ali, and A. Rizki Rinaldi, "Penerapan Metode K-Nearest Neighbor Untuk Prediksi Penjualan Sepeda Motor Terlaris," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 7, no. 1, pp. 585–590, 2023, doi: 10.36040/jati.v7i1.6419.
- [8] T. S. Widiatoro, R. A. Ramadhani, and D. Swanjaya, "Pendekatan Knn Regression Untuk Prediksi Saham Game Stop," *SEMNAS INOTEK(Seminar Nas. Inov. Teknol.)*, vol. 9, pp. 1373–1381, 2025.
- [9] A. Suarisman, A. Nazir, F. Syafria, and L. Afriyanti, "Perbandingan Jarak Metrik pada Klasifikasi Jamur Beracun Menggunakan Algoritma K-Nearest Neighbor (K-NN)," *J. Comput. Syst. Informatics*, vol. 5, no. 1, pp. 10–19, 2023, doi: 10.47065/josyc.v5i1.4511.
- [10] M. G. R. Miftahus Sholihin, "Klasifikasi Mutu Telur Berdasarkan Fitur Warna dengan Menggunakan Metode K-Nearest Neighbor.pdf," *Semnas Sist. Inf.*, p. 6, 2018.
- [11] N. H. Pajriati, E. Kurniati, and D. Suhaedi, "Penerapan Metode Average Based Fuzzy Time Series Lee Untuk Peramalan Harga Emas Di PT. X," *J. Ris. Mat.*, vol. 1, no. 1, pp. 73–81, 2021, doi: 10.29313/jrm.v1i1.221.
- [12] O. J. Ababil, S. A. Wibowo, and H. Zulfia Zahro', "Penerapan Metode Regresi Linier Dalam Prediksi Penjualan Liquid Vape Di Toko Vapor Pandaan Berbasis Website," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 6, no. 1, pp. 186–195, 2022, doi: 10.36040/jati.v6i1.4537.
- [13] P. D. Sugiyono, *Penelitian Metode Kuantitatif, Kualitatif, dan R&D*, Kedua. Bandung: Alfabeta, 2020.
- [14] M. M. Hasmawati, Jumadil Nangi, "Aplikasi Prediksi Penjualan Barang Menggunakan Metode K-Nearest Neighbor (Knn)," vol. 3, no. 2, pp. 151–160, 2017.
- [15] Mustakim and G. Oktaviani F, "Algoritma K-Nearest Neighbor Classification Sebagai Sistem Prediksi Predikat Prestasi Mahasiswa," vol. 13, no. 2, pp. 195–202, 2016.