

# Klasifikasi Kualitas Air Pada Program Penyediaan Air Minum Dan Sanitasi Berbasis Masyarakat Desa Semenpinggir Dengan Metode Algoritma K-Nearest Neighbor

Riwayat Artikel

Received: 1 Jan 2021| Final Revision: 14 feb 2021| Accepted: 14 mar 2021

Mochammad Nur Syarifuddin ✉ #1

# Program studi Teknik Informatika, Nama Universitas Nahdlatul Ulama Sunan Giri Bojonegoro  
Jl. Jendral A. Yani No. 10 Sukorejo Bojonegoro

<sup>1</sup>mochammadnursyarifuddin@gmail.com

**Abstract** — Air bersih merupakan kebutuhan yang sangat penting bagi semua manusia, tersedianya air adalah kebutuhan mutlak untuk menunjang hidup sehat. Penelitian ini memiliki tujuan untuk membuat sebuah aplikasi yang dapat di gunakan mengklasifikasikan kualitas air menggunakan sistem informasi yang dikombinasikan dengan algoritma K-Nearest Neighbor (KNN). Salah satu metode yang di gunakan untuk melakukan klasifikasi adalah Algoritma K-Nearest Neighbor (KNN). Dengan menggunakan data yang dikelompokkan berdasarkan jarak terhadap tetangga terdekat lalu di lakukan perhitungan dengan menggunakan Perhitungan Euclidean untuk kemudian dihitung dan di urutkan yang ada dalam kelompok tersebut. algoritma tersebut diaplikasikan untuk membuat pengelolaan sistem klasifikasi kualitas air yang ada pada Penyediaan Air Minum Dan Sanitasi Berbasis Masyarakat Desa Semenpinggir, Kualitas air yang diuji untuk menghasilkan output berupa klasifikasi menjadi memenuhi syarat dan tidak memenuhi syarat menggunakan parameter berupa Suhu, PH, TDS, Warna, Rasa dan Bau Berdasarkan hasil pengujian diperoleh. Dari hasil perhitungan dengan algoritma K-Nearest Neighbor menggunakan 27 record data, berupa 19 data training dan 8 data testing diperoleh 6 data yang di prediksi memenuhi syarat dan 2 data tidak memenuhi syarat dengan perhitungan menggunakan tetangga terdekat  $K = 3$ . System identifikasi mendapatkan tingkat akurasi 80% berdasarkan perhitungan Confusion Matrix.

**Keywords**— Klasifikasi; K-Nearest Neighbor; Kualitas Air ; Uji Kualitas.

**Abstract** — Clean water is a very important need for all humans, the availability of water is an absolute necessity to support a healthy life. This study aims to create an application that can be used to classify water quality using an information system combined with the K-Nearest Neighbor (KNN) algorithm. One of the methods used to classify is the K-Nearest Neighbor (KNN) Algorithm. By using data that is grouped based on the distance to the nearest neighbor, then the calculation is carried out using Euclidean calculations to then be calculated and sorted in the group. The algorithm was applied to manage the existing water quality classification system in the Provision of Community-Based Drinking Water and Sanitation in Semenpinggir Village. Taste and Odor Based on the test results obtained. From the results of calculations with the K-Nearest Neighbor algorithm using 27 data records, in the form of 19 training data and 8 testing data, 6 data are predicted to meet the requirements and 2 data do not meet the requirements by calculating using the nearest neighbor  $K = 3$ . The identification system gets a level of accuracy 80% based on the Confusion Matrix calculation.

**Keywords**— Classification, K-Nearest Neighbor, Water Quality, Water Quality Test.

## I. PENDAHULUAN

Kebutuhan manusia yang sangat penting salah satunya merupakan air bersih. Karena hampir seluruh aktivitas masyarakat dalam lingkup kehidupan manapun memerlukan ketersediaan air bersih yang cukup. Tersedianya kebutuhan air bersih yang cukup adalah salah satu hal yang mutlak untuk menunjang hidup yang sehat. ketersediaannya sangat penting guna memenuhi kebutuhan yang tidak hanya di

lingkup rumah tangga, akan tetapi juga untuk fasilitas umum, dan juga pertanian guna menunjang ekonomi masyarakat. Agar dapat dimanfaatkan secara baik oleh masyarakat maupun makhluk hidup lainnya itu air harus di lindungi dan di jaga kualitasnya. Pemanfaatan air yang begitu luasnya harus diperhitungkan aspek guna memperhitungkan kepentingan generasi sekarang dan mendatang [1].

Menurut standar air minum Indonesia yang ditetapkan oleh PP No. 82 Tahun 2001 dan KepMen No. 907 Tahun

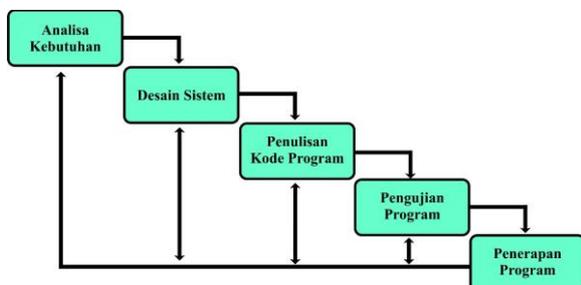
2002, air bersih yang digunakan setiap hari harus berkualitas baik untuk dikonsumsi. Air bersih sangat penting bagi kehidupan manusia sehingga dapat digunakan jika penggunaannya tidak diatur dengan baik maka akan timbul kurangnya sumber air yang layak karena ketersediannya terbatas. Oleh karena itu perlu adanya pengawasan dan control kualitas air yang baik untuk menjaga semua pengguna air tersebut.

Desa Semenpinggir Kecamatan Kapas Kabupaten Bojonegoro adalah satu desa yang cara untuk memperoleh air bersih selain menggunakan PDAM (Perusahaan air minum daerah) adalah dengan memanfaatkan pelayanan PAMSIMAS (Penyediaan Air Minum Sanitasi Berbasis Masyarakat) yang di bangun dan di dirikan pada 2018 melalui bantuan keuangan APBDProv Jawa timur yang pengelolaannya di kelola oleh BUMDes Semenpinggir secara mandiri. Banyaknya resiko dan kebutuhan dalam pemenuhan pelayanan air tersebut sehingga untuk mencapai melaksanakan air bersih yang merata sangatlah tidak mudah, hal ini dikarenakan. Resiko ini dapat bersifat teknis dan non teknis. Misalnya masalah teknis kerusakan alat, kotornya tendon dll yang dapat mempengaruhi standart kelayakan air untuk dapat di konsumsi dan dimanfaatkan.

Terdapat dua tujuan dari penelitian yang ingin di sampaikan yaitu untuk menciptakan sebuah rancangan system yang dapat di dimanfaatkan untuk mengetahui kualitas air. Yang kedua untuk dapat melaksanakan uji kualitas air dan mengetahui tingkat akurasi system identifikasi pengelolaan kualitas air dengan metode Algoritma K-Nearest Neighbor. Kedua tujuan tersebut di diharapkan dapat mengatasi masalah yang selama ini dihadapi pengelola untuk memudahkan melaksanakan uji kualitas air dan dapat mengurangi biaya yang dapat di keluarkan untuk melaksanakan uji lap air. Disamping itu, untuk menerapkan sebuah Algoritma K-Nearest Neighbor yang dapat di gunakan untuk memodelkan kasus prediksi dengan akurasi yang tinggi.

## II. METODE PENELITIAN

Metode penelitian yang digunakan dalam penelitian ini adalah *waterfall* yang dapat dilihat pada Gambar 1 berikut[2]:



Gambar 1. Metodologi Penelitian

### A. Karakteristik Air

Air merupakan salah satu dari sekian banyak kebutuhan manusia di alam. Namun ketersediaan air untuk memenuhi kebutuhan manusia relatif sedikit karena dibatasi oleh

berbagai faktor. Lebih dari 97% air di bumi adalah air laut yang tidak dapat dimanfaatkan secara langsung oleh manusia. Dari 3% dari sisa air, 2% di antaranya disimpan sebagai gunung es kutub (gletser) dan uap air, dan tidak dapat digunakan secara langsung. Hanya 0,62% air yang sebenarnya tersedia untuk kebutuhan manusia, termasuk air yang terdapat di danau, sungai, dan air tanah. Dari segi kualitas, air yang cukup untuk konsumsi manusia hanya menyumbang 0,003% dari seluruh air yang tersedia.

Gambar 2. Persentase ketersediaan air tawar di bumi dengan kualitas yang memadai bagi manusia



Karena perbedaan curah hujan tahunan, air tawar yang dapat dikonsumsi tidak terdistribusi secara merata. Daerah kaya air berada di daerah tropis dan musiman atau beriklim sedang, sedangkan daerah yang kekurangan air berada di daerah kering dan semi-kering. [3].

### B. KNN

Metode K-Nearest Neighbor atau KNN merupakan salah satu metode yang digunakan dalam sistem klasifikasi menggunakan metode machine learning. Pembelajaran mesin adalah proses membangun sistem klasifikasi dengan belajar dari banyak contoh yang sudah diklasifikasikan sebelumnya [2].

Data pelatihan dijelaskan oleh atribut n. Setiap data merupakan titik dalam ruang-n dimensi. Dengan cara ini, semua data pelatihan disimpan di ruang pola-n dimensi. Ketika diberi data yang tidak diketahui, pengklasifikasian k-nearest neighbor akan mencari ruang pola untuk data k pelatihan yang paling dekat dengan data yang tidak diketahui. Data k pelatihan ini adalah k "nearest neighbor" dari data yang tidak diketahui. Dekat atau jauhnya tetangga jarak biasanya dihitung berdasarkan jarak euclidean dengan rumus pada persamaan dibawah. [4].

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

Keterangan :

- $dist$  : jarak
- $x_1$  : nilai data pakar
- $x_2$  : nilai data testing
- $n$  : dimensi data
- $i$  : variabel data

## III. HASIL DAN PEMBAHASAN

Menurut metode penelitian pada Bab II, beberapa hal yang perlu dilakukan untuk menyelesaikan penelitian ini antara lain analisis kebutuhan yang terdiri dari perencanaan, pengumpulan dan pengolahan data, hasil dan analisis, dan dokumentasi.

### A. Pengumpulan Data

Dalam penelitian ini objek penelitian yang dimaksud adalah Kualitas air pada PAMSIMAS Desa Semenpinggir. Dalam penelitian ini data yang digunakan menggunakan data primer dan data sekunder. Data primer yang di ambil berupa pengukuran di lapangan berdasarkan kondisi fisik, kimia dan biologi dari PAMSIMAS Desa Semenpinggir. Data sekunder diperoleh dari beberapa sumber seperti hasil laporan, ataupun dokumen pendukung lain yang berhubungan dengan topik yang diteliti. Data sekunder berupa :

1. Profil PAMSIMAS Desa Semenpinggir
2. Hasil Dari Pengukuran Kualitas Air Dari PAMSIMAS Desa Semenpinggir Sebelumnya.
3. Hasil Dari Pengukuran Kualitas Air Dari PAMSIMAS atau penyedia air lain yang berhubungan dengan data yang di kaji.

Penganalisis klasifikasi data dalam penelitian ini menggunakan algoritma klasifikasi K-nearest neighbor. Prinsip kerja algoritma KNN ditentukan berdasarkan jarak terdekat antara data latihan dan data uji, kemudian mengambil mayoritas sebagai data yang diperoleh digunakan sebagai prediksi untuk pengujian.

Langkah-langkah perhitungan algoritma K-nearest neighbour sebagai berikut[4] :

- a) Tentukan parameter K (jumlah tetangga terdekat)
- b) Hitung kuadrat jarak Euclidean (*query instance*) dari setiap objek ke data sampel yang diberikan.
- c) Kemudian mengurutkan data-data tersebut ke tersebut ke dalam kelompok dengan jarak Euclidean terkecil.
- d) Mengumpulkan kategori Y (Klasifikasi Nearest Neighbor).
- e) Dengan menggunakan kategori Nearest Neighbor sebagai mayoritas tetangga terdekat, maka nilai query instance dihitung dapat diprediksi.

Berikut pada Tabel menjelaskan variabel respon yang digunakan dalam penelitian ini.

Tabel 1 Tabel Variabel

| No | Kode  | Parameter   | Satuan |
|----|-------|---|--------|
| 1  | $x^1$ | Suhu  | °C     |
| 2  | $x^2$ | PH  | -      |
| 3  | $x^3$ | TDS   | mg/l   |
| 4  | $x^4$ | Warna   | -      |
| 5  | $x^5$ | Rasa  | -      |
| 6  | $x^6$ | Bau   | -      |
| 7  | y     | Status Kualitas dari hasil pemeriksaan sample air<br>1 = MS (Memenuhi Syarat),<br>2 = TMS (Tidak Memenuhi Syarat) | -      |

Data nilai nominal akan diubah menjadi data numerik. Pada Tabel 1, data nominal yang masih peril di ubah adalah bau dan rasa. Dalam tabel 2 di bawah adalah hasil perubahan data nominal ke numerik.

Tabel 2 Tabel atribut data

| Atribut        | Nilai |
|----------------|-------|
| Berwarna       | 100   |
| Tidak Berwarna | 200   |
| Berrasa        | 100   |
| Tak berrasa    | 200   |
| Berbau         | 100   |
| Tak berbau     | 200   |

Data yang dilakukan pada bab ini adalah 27 data. Dari data tersebut diambil 19 data yang dijadikan data Training dan 8 data sebagai data Testing. Dari data Training dan data Testing tersebut hasilnya dapat dilihat pada tabel 3 dan 4 sebagai berikut.

Tabel 3 Data Training Kualitas Air Bersih

| No | Tanggal    | Suhu (°C) | PH  | TDS | Warna | Rasa | Bau | Klasifikasi |
|----|------------|-----------|-----|-----|-------|------|-----|-------------|
| 1  | 2/12/2018  | 28        | 7.7 | 620 | 100   | 200  | 200 | TMS         |
| 2  | 2/12/2018  | 30        | 7.3 | 520 | 100   | 100  | 100 | MS          |
| 3  | 2/12/2018  | 29        | 7.5 | 586 | 200   | 200  | 200 | TMS         |
| 4  | 2/12/2018  | 30        | 7.2 | 484 | 100   | 100  | 100 | MS          |
| 5  | 2/12/2018  | 30        | 7.3 | 594 | 200   | 100  | 200 | TMS         |
| 6  | 2/12/2018  | 29        | 7.3 | 455 | 100   | 100  | 100 | MS          |
| 7  | 12/20/2018 | 29        | 7.2 | 502 | 100   | 100  | 100 | MS          |
| 8  | 12/20/2018 | 30        | 7.2 | 489 | 100   | 100  | 100 | MS          |
| 9  | 12/20/2018 | 31        | 7.3 | 510 | 100   | 100  | 100 | MS          |
| 10 | 3/15/2019  | 31        | 7.1 | 513 | 100   | 100  | 100 | MS          |
| 11 | 6/4/2019   | 30        | 7.2 | 495 | 100   | 100  | 100 | MS          |
| 12 | 9/25/2019  | 31        | 7.3 | 480 | 100   | 100  | 100 | MS          |
| 13 | 12/20/2019 | 30        | 7.1 | 504 | 100   | 100  | 100 | MS          |
| 14 | 3/10/2020  | 29        | 7.3 | 493 | 100   | 100  | 100 | MS          |
| 15 | 6/13/2020  | 31        | 7.2 | 479 | 100   | 100  | 100 | MS          |
| 16 | 9/22/2020  | 30        | 7.2 | 507 | 100   | 100  | 100 | MS          |
| 17 | 12/20/2020 | 31        | 7.1 | 512 | 100   | 100  | 100 | MS          |
| 18 | 3/4/2021   | 31        | 7.2 | 498 | 100   | 100  | 100 | MS          |
| 19 | 6/14/2021  | 30        | 7.1 | 507 | 100   | 100  | 100 | MS          |

Tabel 4 Data Testing Kualitas Air Bersih

| No | Tanggal   | Lokasi   | Suhu (°C) | PH  | TDS | Warna | Rasa | Bau | Klasifikasi |
|----|-----------|----------|-----------|-----|-----|-------|------|-----|-------------|
| 1  | 4/08/2021 | PAM Desa | 30        | 7.3 | 198 | 100   | 100  | 100 | ?           |
| 2  | 4/08/2021 | Sampel 1 | 28        | 7.1 | 523 | 100   | 100  | 100 | ?           |
| 3  | 5/08/2021 | Sampel 2 | 29        | 7.3 | 215 | 100   | 100  | 100 | ?           |
| 4  | 5/08/2021 | Sampel 3 | 30        | 7.4 | 593 | 200   | 100  | 200 | ?           |
| 5  | 6/08/2021 | Sampel 4 | 30        | 7.3 | 560 | 200   | 200  | 100 | ?           |
| 6  | 6/08/2021 | Sampel 5 | 31        | 7.1 | 503 | 100   | 200  | 100 | ?           |
| 7  | 7/08/2021 | Sampel 6 | 30        | 7.2 | 488 | 100   | 100  | 100 | ?           |
| 8  | 7/08/2021 | Sampel 7 | 29        | 7.2 | 320 | 100   | 100  | 100 | ?           |

### B. Perhitungan

Perhitungan Euclidean [5]

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

Keterangan :

$dist$  : jarak

$x_1$  : nilai data pakar

$x_2$  : nilai data testing

$n$  : dimensi data

$i$  : variabel data

Dari rumus perhitungan tersebut di sederhanakan dalam bentuk tabel.



|   |           |          |    |     |     |     |     |     |                 |
|---|-----------|----------|----|-----|-----|-----|-----|-----|-----------------|
| 6 | 6/08/2021 | Sampel 5 | 31 | 7.1 | 503 | 100 | 200 | 100 | Memenuhi Syarat |
| 7 | 7/08/2021 | Sampel 6 | 30 | 7.2 | 488 | 100 | 100 | 100 | Memenuhi Syarat |
| 8 | 7/08/2021 | Sampel 7 | 29 | 7.2 | 320 | 100 | 100 | 100 | Memenuhi Syarat |

**D. Perhitungan Confusion Matrix**

Dari data yang telah di Klasifikasi dalam tabel 7 Selanjutnya di lakukan penghitungan Confusion Matrix data yang di peroleh dan dapat di lihat dalam tabel 8. Tabel 8 Tabel penghitungan Confusion Matrix dari Tabel 7.

| Klasifikasi Kebenaran  | Jumlah |
|--|--------|
| True Positive (TP) (Memenuhi Syarat - Memenuhi Syarat)             | 6      |
| True Negative (TN) (Tidak Memenuhi Syarat - Tidak Memenuhi Syarat) | 0      |
| False Positive (FP) (Memenuhi Syarat - Tidak Memenuhi Syarat)      | 0      |
| False Negatif (FN) (Tidak Memenuhi Syarat - Memenuhi Syarat)       | 2      |

Dari Tabel 8 tersebut akan dihitung untuk menentukan nilai dari *precision*, *recall* dan *accuracy*.

Berikut perhitungannya :

$$Precision = \frac{TP}{TP+FP} : \frac{6}{6+0} : 1 : 100\%$$

$$Recall = \frac{TP}{TP+FN} : \frac{6}{6+2} : 0,8 : 80\%$$

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} : \frac{6+0}{6+0+0+2} : 0,8 : 80\%$$

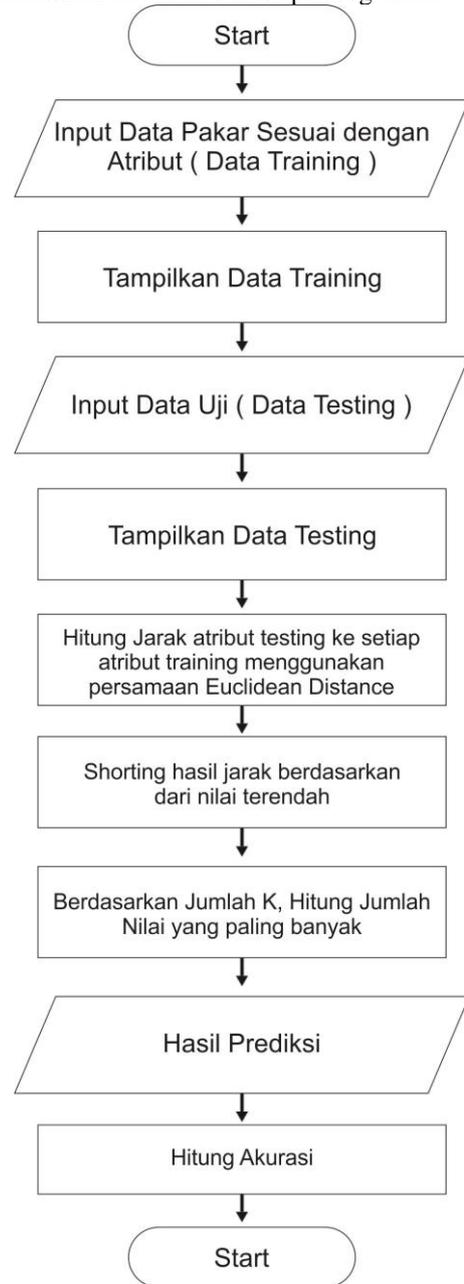
Sehingga di peroleh nilai Precision 100%, Recall 80%, Accuracy 80%. Di presentasikan dalam sebuah diagram sebagai berikut di bawah ini :



Gambar 2 Hasil Akurasi Perhitungan KNN

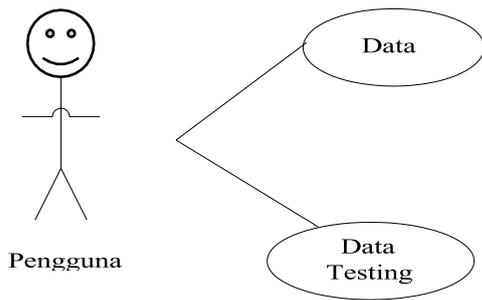
**E. Implementasi Sistem**

Perancangan dan implementasi sistem bertujuan untuk mewujudkan dan mengimplementasikan algoritma yang akan digunakan dan diimplementasikan yang telah dibahas sebelumnya, untuk dapat digunakan dan selanjutnya diimplementasikan kedalam bahasa pemrograman.



Gambar 3 Desain Alur ( Flowchart )

Gambaran *use case* diagram sistem dapat dilihat pada gambar dibawah :



Gambar 4 Gambaran Use case

| No | Total Perhitungan | BK | Klasifikasi     |
|----|-------------------|----|-----------------|
| 10 | 1.00487621121     |    | Memenuhi Syarat |
| 10 | 1.00487621121     |    | Memenuhi Syarat |
| 17 | 3.103358409113    |    | Memenuhi Syarat |

|              |       |
|--------------|-------|
| Akurasi K=3  | 100 % |
| Akurasi K=5  | 100 % |
| Akurasi K=7  | 100 % |
| Akurasi K=9  | 100 % |
| Akurasi K=11 | 100 % |

Gambar 8 Halaman Klasifikasi

| No | Tanggal    | Nama Instansi | Suhu (°C) | PH  | TDS | Warna | Rasa | Bau | Klasifikasi           | Aksi |
|----|------------|---------------|-----------|-----|-----|-------|------|-----|-----------------------|------|
| 1  | 2018-02-12 | PAM Desa      | 28        | 7.3 | 140 | 1     | 2    | 2   | Tidak Memenuhi Syarat | ✖    |
| 2  | 2018-02-12 | PAM Desa      | 30        | 7   | 120 | 1     | 1    | 1   | Memenuhi Syarat       | ✔    |
| 3  | 2018-02-12 | PAM Desa      | 29        | 7.3 | 148 | 1     | 2    | 2   | Tidak Memenuhi Syarat | ✖    |
| 4  | 2018-02-12 | PAM Desa      | 30        | 7   | 110 | 1     | 1    | 1   | Memenuhi Syarat       | ✔    |
| 5  | 2018-02-12 | PAM Desa      | 30        | 7.3 | 201 | 2     | 1    | 1   | Tidak Memenuhi Syarat | ✖    |

Gambar 5 Halaman Data Training

| No | Tanggal    | Nama Instansi | Suhu (°C) | PH  | TDS | Warna | Rasa | Bau | Klasifikasi           | Aksi |
|----|------------|---------------|-----------|-----|-----|-------|------|-----|-----------------------|------|
| 1  | 2018-02-12 | PAM Desa      | 28        | 7.3 | 140 | 1     | 2    | 2   | Tidak Memenuhi Syarat | ✖    |
| 2  | 2018-02-12 | PAM Desa      | 30        | 7   | 120 | 1     | 1    | 1   | Memenuhi Syarat       | ✔    |
| 3  | 2018-02-12 | PAM Desa      | 29        | 7.3 | 148 | 1     | 2    | 2   | Tidak Memenuhi Syarat | ✖    |
| 4  | 2018-02-12 | PAM Desa      | 30        | 7   | 110 | 1     | 1    | 1   | Memenuhi Syarat       | ✔    |
| 5  | 2018-02-12 | PAM Desa      | 30        | 7.3 | 201 | 2     | 1    | 1   | Tidak Memenuhi Syarat | ✖    |

Gambar 6 Halaman Data Testing

| No | K1-X | K2-X               | K3-X | K4-X | K5-X | K6-X | Total Perhitungan | Klasifikasi           |
|----|------|--------------------|------|------|------|------|-------------------|-----------------------|
| 1  | 4    | 0.04               | 1300 | 0    | 1    | 1    | 37.28013171466    | Tidak Memenuhi Syarat |
| 2  | 0    | 0.0000000000000000 | 200  | 0    | 0    | 0    | 17.000284191602   | Memenuhi Syarat       |
| 3  | 1    | 0.04               | 2525 | 0    | 1    | 1    | 48.032760191902   | Tidak Memenuhi Syarat |
| 4  | 0    | 0.0000000000000000 | 40   | 0    | 0    | 0    | 7.000716402749    | Memenuhi Syarat       |
| 5  | 0    | 0.04               | 960  | 1    | 0    | 0    | 16.00230479090    | Tidak Memenuhi Syarat |
| 6  | 1    | 0                  | 120  | 0    | 0    | 0    | 21.00100880442    | Memenuhi Syarat       |
| 7  | 1    | 0.0000000000000000 | 120  | 0    | 0    | 0    | 23.0019480154     | Memenuhi Syarat       |
| 8  | 0    | 0.01               | 404  | 0    | 0    | 0    | 24.001008801303   | Memenuhi Syarat       |
| 9  | 1    | 0.04               | 576  | 0    | 0    | 0    | 24.001008801303   | Memenuhi Syarat       |
| 10 | 1    | 0.0000000000000000 | 0    | 0    | 0    | 0    | 1.004887621121    | Memenuhi Syarat       |
| 11 | 0    | 0                  | 144  | 0    | 0    | 0    | 12                | Memenuhi Syarat       |
| 12 | 1    | 0.01               | 720  | 0    | 0    | 0    | 27.000807228881   | Memenuhi Syarat       |
| 13 | 0    | 0.0000000000000000 | 200  | 0    | 0    | 0    | 17.000284191602   | Memenuhi Syarat       |
| 14 | 1    | 0.04               | 220  | 0    | 0    | 0    | 16.002304790902   | Memenuhi Syarat       |
| 15 | 1    | 0.0000000000000000 | 100  | 0    | 0    | 0    | 11.001008801304   | Memenuhi Syarat       |
| 16 | 0    | 0.01               | 360  | 0    | 0    | 0    | 18.002016402749   | Memenuhi Syarat       |

Gambar 7 Halaman Proses Prediksi

#### IV. SIMPULAN

Berdasarkan hasil penelitian peneliti terhadap analisis dan implementasi sistem pengenalan kualitas air menggunakan algoritma K-nearest neighbor dapat ditarik beberapa kesimpulan, yaitu: Penelitian ini berhasil menerapkan algoritma K-nearest neighbor untuk menghitung dan memberikan hasil klasifikasi kualitas air. Penelitian ini menggunakan 27 data record, 19 data sebagai data latih, dan 8 data sebagai data uji. Dari data testing tersebut 6 data diprediksi benar dan 2 data diprediksi salah dengan perhitungan menggunakan nilai K = 3.

Sistem untuk identifikasi kualitas air dengan menggunakan perhitungan Confusion Matrix untuk metode KNN hasilnya akurat dengan mencapai tingkat akurasi sebesar 80 %.

#### DAFTAR PUSTAKA

- [1] F. Liantoni, "Klasifikasi Daun Dengan Perbaikan Fitur Citra Menggunakan Metode K-Nearest Neighbor," *J. Ultim.*, vol. 7, no. 2, pp. 98–104, 2016.
- [2] G. W. Sasmito, "Penerapan Metode Waterfall Pada Desain Sistem Informasi Geografis Industri Kabupaten Tegal," *J. Inform. Pengemb. IT*, vol. 2, no. 1, pp. 6–12, 2017.
- [3] S. H. Wardani, T. Rismawan, and S. Bahri, "Aplikasi Klasifikasi Jenis Tumbuhan Mangrove Berdasarkan Karakteristik Morfologi Menggunakan Metode K-Nearest Neighbor (KNN) Berbasis Web," *J. Coding, Sist. Komput. Untan*, vol. 4, no. 3, pp. 9–21, 2016.
- [4] A. A. Aldino, R. R. Suryono, and R. Ambarwati, "Analysis of Covid-19 Cash Direct Aid (BLT) Acceptance Using K-Nearest Neighbor Algorithm," *IJCS (Indonesian J. Comput. Cybern. Syst.)*, vol. 16, no. 2, p. 193, 2022.
- [5] A. P. Pawlovsky, "An ensemble based on distances for a kNN method for heart disease diagnosis," *Int. Conf. Electron. Inf. Commun. ICEIC 2018*, vol. 2018-Janua, pp. 1–4, 2018.