

## ANALISIS KEPUASAN PELANGGAN MENGGUNAKAN ALGORITMA K-NEAREST NEIGHBORS PADA PERUSAHAAN UMUM DAERAH AIR MINUM BATIWAKKAL BERAU

NOVIA RAHMADANA<sup>1</sup>, ABDUL RAHIM\*<sup>2</sup>,  
FENDY YULIANTO<sup>3</sup>

<sup>1,2,3</sup>Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Muhammadiyah Kalimantan Timur  
email: [ar622@umkt.ac.id](mailto:ar622@umkt.ac.id)

### ABSTRACT

Kepuasan pelanggan terhadap layanan air bersih merupakan aspek penting yang perlu dianalisis oleh Perusahaan Umum Daerah Air Minum Batiwakkal Berau untuk memastikan pelayanan terbaik. Penelitian ini bertujuan mengimplementasikan algoritma K-Nearest Neighbors (KNN) untuk menganalisis data kepuasan pelanggan Air Minum Batiwakkal Berau. Penelitian ini mengevaluasi akurasi model KNN dalam mengklasifikasikan tingkat kepuasan pelanggan, menentukan nilai terbaik parameter K, dan menyediakan gambaran mengenai persepsi pelanggan terhadap layanan Perusahaan Umum Daerah Air Minum Batiwakkal Berau. Metode penelitian melibatkan pengumpulan data dari pelanggan Air Minum Batiwakkal Berau. Data dipisahkan menjadi set pelatihan dan pengujian, diikuti dengan proses normalisasi. Model KNN dilatih dengan berbagai nilai K untuk menemukan parameter terbaik. Hasil penelitian menunjukkan bahwa algoritma KNN dapat mengklasifikasikan tingkat kepuasan pelanggan dengan akurasi memadai. Nilai terbaik K yang ditemukan adalah 14, memberikan performa terbaik dalam prediksi kepuasan pelanggan dengan nilai akurasi sebesar 85.96%.

**Kata Kunci:** Kepuasan pelanggan, Air Minum Batiwakkal Berau, K-Nearest Neighbors

### I. PENDAHULUAN

Air merupakan kebutuhan mendasar yang sangat penting bagi semua bentuk kehidupan di bumi termasuk manusia. Tanpa air, manusia akan mengalami

kesulitan dalam menjalani kehidupan mereka dengan baik, oleh karena itu, pengelolaan dan pengolahan air harus diatur dengan hati-hati untuk menjamin penggunaannya secara efektif dan efisien (Amrulloh, 2022). Dengan bertambahnya jumlah penduduk, permintaan terhadap air bersih juga mengalami peningkatan. Untuk memenuhi kebutuhan terhadap air bersih, pemerintah mengoperasikan sebuah perusahaan umum milik daerah yang disebut Perusahaan Daerah Air Minum (PDAM) (Akhirina et al., 2019). PDAM berperan sebagai salah satu entitas usaha milik daerah yang bertanggung jawab atas penyediaan air bersih dan sehat bagi masyarakat rumah tangga serta perusahaan (Arin Nurwidya, Hery Sawiji, 2022).

Perusahaan Daerah Air Minum (PDAM) Batiwakkal adalah institusi yang berwenang dalam penyediaan air bersih di Kabupaten Berau yang dibangun oleh Pemerintah Pusat pada tahun 1980 (*Sejarah - PERUMDA Air Minum Batiwakkal Kabupaten Berau*, 2022). PDAM Batiwakkal Berau tidak hanya diwajibkan untuk memenuhi kebutuhan air bersih, tetapi juga harus memberikan layanan yang mudah, cepat, akurat bagi pelanggan dan calon pelanggan. Pelayanan pelanggan yang prima adalah salah satu indikator penting keberhasilan PDAM, karena seluruh biaya operasional bergantung pada pelanggan sebagai sumber pendanaan utama bagi kelangsungan operasional perusahaan dalam melayani masyarakat. Oleh karena itu, hubungan serta dukungan masyarakat yang baik harus dibina dan dipelihara melalui PDAM (Haslinda, 2022).

Kepuasan pelanggan membawa dampak positif bagi perusahaan dengan meningkatkan daya tarik dan menjaga loyalitas konsumen. Kepuasan juga menciptakan hubungan harmonis antara perusahaan dan pelanggan, yang penting untuk pembelian berulang dan pembangunan loyalitas pelanggan (Ni Luh Nurkariani, 2022). PDAM Batiwakkal Berau sebagai penyedia air minum bagi masyarakat membutuhkan pemahaman mendalam tentang kepuasan pelanggan. Dengan menerapkan algoritma KNN, PDAM dapat mengidentifikasi pola preferensi dan kepuasan pelanggan, yang dapat digunakan untuk meningkatkan layanan.

Adapun penelitian sebelumnya yang memiliki kesamaan tema dengan

penelitian ini, yaitu: (i) penelitian yang dilakukan oleh Dandi Junandi dan Muhammad Ridwan. Dari hasil penelitian yang dilakukan dengan menggunakan data kepuasan pelanggan PDAM Tirtalihou, hasil uji coba yang dilakukan dengan software RapidMiner 5.3 memperoleh nilai akurasi 98.33% (Dandi Junandi Purba, Muhammad Ridwan Lubis, 2022). (ii) penelitian yang dilakukan oleh Faisal dan Nurhayati, pada penelitian ini menggunakan data kuesioner kepuasan pelanggan Cikarang Camera Rental. Data set ini berisi informasi mengenai kuesioner kepuasan pelanggan terkait harga, fasilitas, layanan, dan loyalitas, total data dalam set data ini adalah 100 data. Hasil penelitian ini dengan menggunakan algoritma KNN mendapatkan akurasi= 98%, klasifikasi recall= 86,67%, ketepatan klasifikasi = 100% dan AUC = 0,750 (Faisal & Nurhayati, 2020).

K-Nearest Neighbors adalah salah satu algoritma dalam pembelajaran mesin yang dapat digunakan untuk menganalisis pola-pola dalam data dan membuat prediksi berdasarkan data yang ada (Karomi, 2020). Dalam konteks analisis kepuasan pelanggan, K-Nearest Neighbors dapat membantu mengidentifikasi pola-pola perilaku pelanggan berdasarkan data historis, sehingga memberikan wawasan yang berharga bagi perusahaan (Indrastata, 2023). K-Nearest Neighbors memiliki sejumlah keunggulan, di antaranya pelatihannya yang sangat cepat, sederhana dan mudah dipahami, serta efektif untuk data pelatihan berukuran besar (Muhammad Jauhar Vikri & Rohmah, 2022).

Atas dasar pemikiran inilah yang mengunggah perhatian penulis untuk melaksanakan penelitian yang berjudul "Analisis Kepuasan Pelanggan Menggunakan Algoritma K-Nearest Neighbors Pada Perusahaan Umum Daerah Air Minum Batiwakkal Berau".

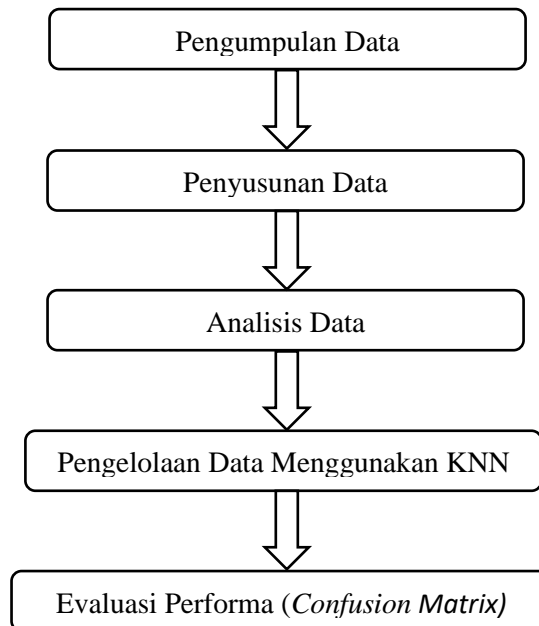
## **II. METODOLOGI PENELITIAN**

### **A. Obyek Penelitian**

Obyek penelitian ini adalah Perusahaan Umum Daerah Air Minum Batiwakkal Berau. Subjek penelitian ini adalah pelanggan yang menggunakan air minum Batiwakkal Berau. Pengumpulan data dilakukan pada tahun 2022 dengan menyebarkan kuisoner kepada pelanggan Air Minum Batiwakkal Berau, data yang digunakan pada penelitian ini yaitu berjumlah 3.572 data.

## B. Prosedur Penelitian

Agar penelitian ini berjalan dengan lancar dan selesai tepat waktu, diperlukan adanya prosedur penelitian. Prosedur penelitian untuk penelitian ini dapat dilihat berdasarkan Gambar 1.



Gambar 1. Prosedur Penelitian

### 1. Melakukan Pengumpulan Data

Pengumpulan data diperlukan untuk mengklasifikasi tingkat kepuasan pelanggan di Air Minum Batiwakkal Berau menggunakan data yang akurat agar kasus pada laporan penelitian ini berkaitan dan sesuai dengan yang dikerjakan penulis. Penulis mengumpulkan data dari kuisisioner dengan teknik survey dimana kuisisioner tersebut disebarakan kepada pelanggan di Air Minum Batiwakkal Berau untuk memperoleh informasi yang lebih mendalam dan spesifik. Pengumpulan data adalah proses mengumpulkan informasi penelitian dari sumber data, yaitu subjek atau sampel penelitian (Halim et al., 2023).

### 2. Analisis Data

Penulis menganalisis data yang telah terkumpul dan data yang diperoleh dari hasil kuisisioner pelanggan Air Minum Batiwakkal Berau.

### 3. Penyusunan Data

Penyusunan data merupakan tahap krusial dalam penelitian di mana data

diklasifikasikan dan disusun secara teratur untuk mempermudah analisis dan interpretasi (Babbie, 2020). Setelah analisis data selesai penulis akan menyusun atau mengumpulkan data yang akan diolah lagi ke tahap selanjutnya. Penyusunan data juga mencakup metode untuk mendeteksi kemungkinan kesalahan, ketidakkonsistenan, ketidakteraturan, serta ketidaktepatan dalam data yang telah dikumpulkan (Maulid, 2021).

#### **4. Pengelolaan Data Menggunakan KNN**

Pengolahan data yang sudah diperoleh akan diolah menggunakan metode KNN untuk melakukan klasifikasi data. *Python* merupakan program gratis, yaitu *freeware* atau program bebas yang dikenal karena distribusinya. Semua fitur terdiri dari kode sumber, debugger, profiler, antarmuka pengguna grafis (GUI) dan data inti (Clinton & Sengkey, 2019). Berikut ini adalah langkah-langkah kerja K-Nearest Neighbors (Laraswati, 2022): (i) tentukan parameter K (banyak tetangga terdekat), (ii) Hitung jarak data baru atau data testing dengan semua data yang berada di data training menggunakan jarak Euclidean, (iii) urutkan jarak dan tentukan tetangga mana yang paling dekat berdasarkan jarak minimum ke-K, (iv) Menentukan kategori dari tetangga terdekat, (v) menggunakan kategori mayoritas yang sederhana dari tetangga terdekat sebagai nilai prediksi data baru.

#### **5. Evaluasi Performa Menggunakan Confusion Matrix**

Confusion matrix adalah tabel yang menyatakan klasifikasi jumlah data uji yang benar dan jumlah data uji yang salah (Evalin et al., 2021).

### **III. HASIL DAN PEMBAHASAN**

#### **A. Data Processing**

##### **1. Cleaning Data.**

Dapat dilihat pada Gambar 2, merupakan hasil dari pembersihan data yang dilakukan pada dataset. Data yang tidak valid telah dihapus atau diperbaiki, data duplikat telah dihilangkan, dan nilai-nilai yang hilang telah diisi menggunakan metode yang sesuai untuk memastikan kualitas data yang optimal sebelum melanjutkan ke tahap analisis lebih lanjut.

	Bagaimana kualitas air yang anda pakai dari pdam?	Apakah jumlah air yang anda perlukan sudah dapat dipenuhi oleh pdam?	Apakah anda dapat menikmati air selama 24 jam dalam sehari?	Secara umum bagaimana pelayanan perumda air minum bativakkal secara umum?
0	3	3	3	4
1	3	3	3	4
2	3	3	3	4
3	3	3	3	4
4	3	3	3	4
...	...	...	...	...
3567	3	2	3	2
3568	3	3	4	4

Gambar 2. Hasil dari pembersihan data

## 2. Memisahkan Fitur (Features) Dan Label (Target) Dalam Dataset Dan Normalisasi Data

Pada Gambar 3 normalisasi data dilakukan untuk mengubah skala atau rentang nilai ke dalam rentang yang lebih terstandarisasi atau normal, agar tidak ada fitur yang mendominasi yang lain dalam perhitungan jarak pada algoritma KNN.

```
[9] #memisahkan fitur dan label
x = selected_dataset.iloc[:, :-1].values
y = selected_dataset.iloc[:, -1].values

#normalisasi data
scaler = MinMaxScaler()
x = scaler.fit_transform(x)
```

Gambar 3. Code memisahkan fitur dan label dan normalisasi data

## B. Pengelolaan Data Menggunakan KNN

### 1. Memisahkan Data *Train* dan Data *Test*

Pada Gambar 4 dengan membagi data menjadi data pelatihan (*train*) dan data pengujian (*test*), untuk dapat menguji performa model pada data yang tidak pernah dilihat sebelumnya.

```
[10] #memisahkan data train dan data test
x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=0.2, random_state=42)
```

Gambar 4. Memisahkan data train dan test

### 2. Mencari Nilai K Yang Optimal

Pada Gambar 5 ditemukan bahwa panjang *k\_range* dan *mse* adalah 30. Setelah memilih nilai K yang memiliki *mean error* terkecil, yaitu 14, model KNN dianggap terbaik untuk *dataset* yang digunakan. Visualisasi dari *plot Error vs Nilai K* juga menggambarkan bahwa *mean squared error* cenderung menurun dengan peningkatan nilai k hingga mencapai titik terendah pada nilai K=14, sebelum sedikit naik kembali.

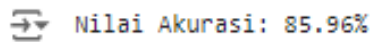




**C. Evaluasi Performa Menggunakan Confusion Matrix**

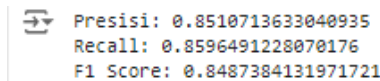
**1. Menghitung Akurasi, Precision, Recall dan F-1 Score**

Pada Gambar 8 kode ini bertujuan untuk menghitung dan mencetak nilai akurasi dari model *K-Nearest Neighbors* yang telah dilatih dan diuji terhadap data uji, mendapatkan nilai akurasi sebesar 85.96%.



Gambar 8. Hasil akurasi menggunakan KNN

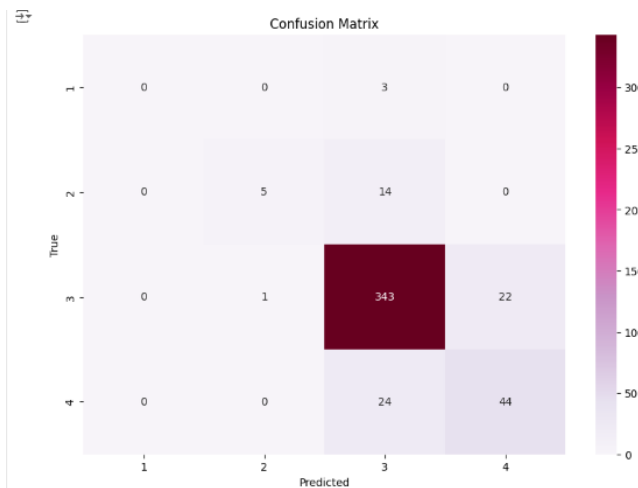
Pada Gambar 9, dalam hasil evaluasi yang telah dilakukan, nilai *precision*, *recall*, dan *F1-Score* adalah *Precision: 0.8517*, *Recall: 0.8596*, *F1-Score: 0.848*.



Gambar 9. Menghitung precision, recall, dan f1-score

**2. Menampilkan Visualisasi Confusion Matrix Dalam Bentuk Grafik Heatmap**

Pada Gambar 10 merupakan hasil dari memvisualisasikan *confusion matrix* dari model *K-Nearest Neighbors* yang telah dilatih dan diuji terhadap data uji.



Gambar 10. Menampilkan visualisasi confusion matrix dalam bentuk grafik heatmap

Untuk membandingkan hasil evaluasi tersebut maka dilakukan perhitungan manual nilai *confusion matrix* untuk mengkonfirmasi keakuratan hasil evaluasi.



**Tabel 1.** Perhitungan manual

Hasil perhitungan manual	
Akurasi	0.8593
Presision	0.8615
Recall	0.8596
F1 Score	0.8655

Setelah melakukan perhitungan manual pada setiap parameter, hasilnya sama dengan perhitungan menggunakan *python*. Ini menunjukkan kesesuaian yang sangat baik dengan perhitungan manual, yang mengindikasikan bahwa model memiliki kemampuan kuat dalam menerapkan data latih ke data uji dan mampu mengklasifikasikan data dengan akurat.

#### IV. KESIMPULAN

Kesimpulan dari penelitian ini menunjukkan bahwa algoritma *K-Nearest Neighbors* (KNN) dapat secara efektif digunakan untuk menganalisis kepuasan pelanggan terhadap layanan PDAM. Melalui proses pemisahan data, normalisasi, dan pemilihan nilai K yang terbaik, model KNN berhasil mengklasifikasikan tingkat kepuasan pelanggan dengan akurasi yang memadai. Hasil analisis menunjukkan bahwa nilai terbaik K adalah 14 dan mendapatkan nilai akurasi sebesar 85.96% yang memberikan performa terbaik dalam memprediksi kepuasan pelanggan berdasarkan fitur-fitur yang ada.

#### DAFTAR PUSTAKA

- Akhirina, C., Gumilar, N., Sianturi, C., Surya, R., Wistyani, M., Baknur, F., Kristina, Y., & Larasati, E. (2019). *PDAM Sehat dan Mandiri Melalui Kerjasama Investasi*. 31–35.
- Amrulloh, R. (2022). Survei Kepuasan Masyarakat Terhadap Kualitas Pelayanan Pdam Intan Banjar. *Al-Kalam : Jurnal Komunikasi, Bisnis Dan Manajemen*, 9(2), 128. <https://doi.org/10.31602/al-kalam.v9i2.6526>
- Arin Nurwidya, Hery Sawiji, T. M. (2022). Analisis Kepuasan Pelanggan Pada Pelayanan Air Bersih Di PDAM Tirta Bening Kabupaten Pati. *Jurnal Informasi Dan Komunikasi Administrasi Perkantoran Volume*, 6(2), 112.
- Babbie, E. R. (2020). *The Practice of Social Research*. Cengage Learning. [https://books.google.co.id/books?hl=id&lr=&id=IFvjDwAAQBAJ&oi=fnd&pg=PP1&dq=Babbie,+E.+R.+\(2020\).+The+Practice+of+Social+Research.+C](https://books.google.co.id/books?hl=id&lr=&id=IFvjDwAAQBAJ&oi=fnd&pg=PP1&dq=Babbie,+E.+R.+(2020).+The+Practice+of+Social+Research.+C)

- engage+Learning.&ots=I4xY\_G3TP9&sig=T0kdjWmBHp5vioUc5A-ZnEE5NE&redir\_esc=y#v=onepage&q=Babbie%2C E. R. (2020). The Practic
- Clinton, R. M. R., & Sengkey, S. (2019). Purwarupa Sistem Daftar Pelanggaran Lalulintas. *Jurnal Teknik Elektro Dan Komputer Vol.8*, 8(3), 181–192.
- Dandi Junandi Purba, Muhammad Ridwan Lubis, Z. A. S. (2022). Analisis Kepuasan Pelanggan Menggunakan Algoritma C4.5. *Jurnal Teknik Informatika UNIKA Santo Thomas*, 2(2), 115–121. <https://doi.org/10.54367/jtiust.v7i1.1983>
- Evalin, Z. Z., Sukartaatmadja, I., & Mulyadi, M. N. (2021). Analisis Peran Kinerja Customer Service Dan Teller Terhadap Kepuasan Nasabah Bank. *Jurnal Ilmiah Manajemen Kesatuan*, 9(2), 151–160. <https://doi.org/10.37641/jimkes.v9i2.767>
- Faisal, S., & Nurhayati. (2020). Implementation of K-Nearest Neighbor Algorithm for Customer Satisfaction. *Buana Information Technology and Computer Sciences (BIT and CS)*, 1(2), 27–32. <https://doi.org/10.36805/bit-cs.v1i2.886>
- Halim, T. N., Martin, R., & ... (2023). Klasifikasi Kepuasan Pelanggan Terhadap Platform E-Commerce dengan Metode K–Nearest Neighbor (K-NN). *Jurasik (Jurnal Riset ...)*, 8, 512–523. <http://ejurnal.tunasbangsa.ac.id/index.php/jurasik/article/view/636%0Ahttps://ejurnal.tunasbangsa.ac.id/index.php/jurasik/article/download/636/609>
- Haslinda. (2022). *Analisis Sistem Informasi Akutansi Penerimaan Kas PERUMD Air Minum Batiwakkal Berau*.
- Indrastata, I. (2023). Strategi Promosi Perbankan Dengan Analisis Dan Implementasi Algoritma Random Forest. *AnoaTIK: Jurnal Teknologi Informasi Dan Komputer*, 1(2), 54–58. <https://doi.org/10.33772/anoatik.v1i2.8>
- Karomi, M. A. Al. (2020). Optimasi Parameter K Pada Algoritma Knn Untuk Klasifikasi Heregistrasi Mahasiswa. *IC-Tech*, 10(1), 28–33.
- Laraswati, B. D. (2022). *Mengenal K-nearest Neighbor dan Pengaplikasiannya*. Algoritma. <https://blog.algorit.ma/k-nearest-neighbor/#cara-kerja-k-nearest-neighbor>
- Maulid, R. (2021). *Simak Langkah-Langkah dalam Teknik Pengolahan Data*. <https://dqlab.id/simak-langkah-langkah-dalam-teknik-pengolahan-data>
- Muhammad Jauhar Vikri, & Rohmah, R. (2022). Penerapan Fungsi Exponential Pada Pembobotan Fungsi Jarak Euclidean Algoritma K-Nearest Neighbor. *Generation Journal*, 6(2), 57–64. <https://doi.org/10.29407/gj.v6i2.18070>
- Ni Luh Nurkariani, G. Y. K. (2022). Pentingnya Kepuasan Konsumen Untuk Meningkatkan Loyalitas Pelanggan. *Artha Satya Dharma*, 15(1), 27–32. <https://doi.org/10.55822/asd.v15i1.238>
- Sejarah - PERUMDA Air Minum Batiwakkal Kabupaten Berau*. (2022). <https://perumdabatiwakkal.co.id/profil/sejarah/>