

POLA KEPUASAN PELANGGAN PDAM MENGGUNAKAN DATA MINING PADA BIG DATA OPERASIONAL DAN LAYANAN

Nanang Fahrurrozi*, Joko Handoyo

Program Studi Informatika, Sekolah Tinggi Teknologi Ronggolawe Cepu.

*kamajatifahru@gmail.com

Jl. Kampus Ronggolawe Blok B No.1 Mentul, Blora, Jawa Tengah, Indonesia.

Article history: Received: 15 March 2026; Revised: 25 May 2026; Accepted: 30 June 2026

Abstract

This research addresses the challenge faced by PDAM in holistically understanding customer satisfaction due to fragmented data. By applying data mining to integrated big data from Perumda Air Minum Tirta Amerta Blora, the study combines 7,766 data points from technical reports, meter calibration records, SIMPel complaints, the Si Tampan application, and social media using the CRISP-DM methodology. K-Means Clustering analysis identified four customer segments. Loyal Satisfied at 18%, Technical Complaints at 32%, Administrative Issues at 29%, and dissatisfied at 21%. Three sub-districts—Cepu, Ngawen, and Randublatung—were confirmed as critical areas with a Silhouette Score of 0.58. Random Forest Classification revealed five dominant factors influencing satisfaction. complaint response time at 0.224, water pressure at 0.198, disruption frequency at 0.156, meter status at 0.142, and arrears at 0.134. A strong correlation was found between water pressure and complaint frequency, with $r = -0.72$. The predictive model demonstrated excellent performance, achieving an accuracy of 89%, precision of 88%, recall of 85%, F1-score of 0.86, and AUC-ROC of 0.93. TF-IDF sentiment analysis confirmed the dominance of technical complaints through words such as "leak" and "no water flow," as well as appreciation for digital innovation, with 54% positive sentiment on social media. These findings recommend prioritizing infrastructure improvements in critical areas, optimizing complaint response time to less than 8 hours, developing a real-time analytics dashboard, strengthening the Si Tampan digital service, and implementing a proactive social media strategy with rapid response to sustainably enhance customer satisfaction.

Keywords: *Customer Satisfaction; Data Mining; K-Means Clustering; Random Forest; Sentiment Analysis.*

Abstrak

Penelitian ini mengatasi tantangan PDAM dalam memahami kepuasan pelanggan secara holistik akibat data yang terfragmentasi. Dengan menerapkan data mining pada big data terintegrasi Perumda Air Minum Tirta Amerta Blora, penelitian menggabungkan 7.766 data dari laporan teknik, kalibrasi meter, pengaduan SIMPel, aplikasi Si Tampan, dan media sosial menggunakan metodologi CRISP-DM. Analisis K-Means Clustering mengidentifikasi empat segmen pelanggan. Loyal Puas sebesar 18%, Keluhan Teknis sebesar 32%, Masalah Admin sebesar 29%, dan Tidak Puas sebesar 21%. Tiga wilayah Cepu, Ngawen, dan Randublatung, dikonfirmasi sebagai area kritis dengan Silhouette Score 0,58. Random Forest Classification mengungkapkan lima faktor dominan kepuasan. waktu respons keluhan sebesar 0,224, tekanan air sebesar 0,198, frekuensi gangguan sebesar 0,156, status meter sebesar 0,142, dan tunggakan sebesar 0,134. Korelasi kuat ditemukan antara tekanan air dan frekuensi keluhan yaitu $r = -0,72$. Model prediktif menunjukkan performa sangat baik dengan akurasi 89%, presisi 88%, recall 85%, F1-score 0,86, dan AUC-ROC 0,93. Analisis sentimen TF-IDF mengonfirmasi dominasi keluhan teknis pada kata "bocor" dan "air tidak keluar", serta apresiasi terhadap inovasi digital dengan 54% sentimen positif di media sosial. Temuan ini merekomendasikan prioritas perbaikan infrastruktur di wilayah kritis, optimalisasi waktu respons



keluhan kurang dari 8 jam, pengembangan *dashboard analitik real-time*, penguatan layanan digital Si Tampan, serta strategi media sosial proaktif dengan respons cepat untuk meningkatkan kepuasan pelanggan secara berkelanjutan.

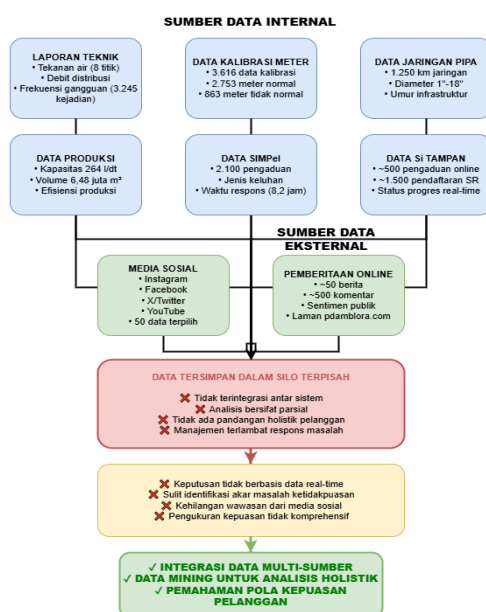
Kata Kunci: Analisis Sentimen; Data Mining; Kepuasan Pelanggan; K-Means Clustering; Random Forest.

Pendahuluan

Perumda Air Minum Tirta Amerta Kabupaten Blora mengemban tanggung jawab strategis dalam penyediaan air bersih untuk kebutuhan masyarakat. Mengacu pada catatan operasional tahun 2025, institusi ini mengelola 25.565 sambungan aktif dengan kapasitas produksi riil sebesar 264 liter per detik, serta total produksi tahunan yang mencapai 6.479.672 m³. Ekspansi layanan terlihat dari penambahan 945 sambungan baru dan perluasan jaringan pipa sepanjang 10.456 meter, mencerminkan komitmen pertumbuhan layanan yang berkelanjutan. Berbagai penelitian terdahulu menggarisbawahi bahwa dimensi pelayanan dan ketersediaan fasilitas memegang peranan penting dalam membentuk tingkat kepuasan konsumen jasa PDAM [1], [2], [3].

Dalam kerangka transformasi digital, perusahaan meluncurkan aplikasi Si Tampan pada Februari 2024 guna memfasilitasi proses pendaftaran serta penanganan keluhan secara daring. Adopsi teknologi serupa di berbagai wilayah lain terbukti memberikan dampak positif terhadap kepuasan pelanggan [4], [5]. Aplikasi ini memperkaya data pengaduan yang sebelumnya dihimpun melalui sistem SIMPel, yang mencatat kurang lebih 2.100 aduan dalam rentang November 2023 hingga Desember 2024. Pada era digital saat ini, pemahaman terhadap opini publik secara langsung juga banyak ditentukan oleh interaksi yang berlangsung di media sosial [6].

Kendati demikian, seluruh data operasional dan layanan tersebut masih terpisah dalam unit-unit penyimpanan yang tidak saling terhubung, sehingga potensinya untuk dianalisis secara mendalam belum terealisasi secara maksimal [7]. Akibat tidak terintegrasinya data laporan teknis, hasil kalibrasi meter, catatan pengaduan dari SIMPel dan Si Tampan, serta percakapan di media sosial, perusahaan kesulitan memperoleh perspektif utuh mengenai kepuasan pelanggan. Padahal, pemahaman komprehensif seperti itu mutlak diperlukan sebagai landasan perumusan kebijakan peningkatan layanan yang lebih akurat dan tepat guna.



Gambar 1. Diagram Konteks Permasalahan

Gambar 1 menyajikan diagram konteks yang memvisualisasikan persoalan fragmentasi data ini, di mana beragam sumber informasi belum tersinergi untuk mendukung proses analisis yang terpadu. Oleh karena itu, langkah mengintegrasikan data dari berbagai sumber menjadi tahap krusial dalam upaya mengidentifikasi pola, faktor determinan, serta sentimen konsumen secara presisi. Dengan demikian, penerapan teknik data mining pada data yang telah terintegrasi diharapkan dapat menghasilkan wawasan strategis guna meningkatkan mutu pelayanan dan kepuasan pelanggan Perumda Air Minum Tirta Amerta Kabupaten Blora secara berkesinambungan.

1. Kepuasan Pelanggan dan Model SERVQUAL

Tingkat kepuasan pelanggan merupakan suatu kondisi psikologis yang diukur seberapa besar kinerja layanan yang diterima mampu memenuhi atau melampaui harapan pelanggan. Kualitas layanan dalam model SERVQUAL didasarkan pada lima indikator utama, meliputi *reliability* (keandalan), *responsiveness* (daya tanggap), *assurance* (jaminan), *emphaty* (empati), dan *tangible* (bukti fisik). Dalam konteks PDAM, dimensi ini diterjemahkan ke dalam indikator teknis seperti kontinuitas aliran, tekanan air, akurasi meter, dan respons keluhan [8], [2], [3]. Prakoso dkk [1] dalam studinya menunjukkan bahwa kualitas pelayanan, kepuasan pelanggan PDAM dipengaruhi oleh fasilitas, dan ketersediaan produk.

2. Big Data dalam Industri Utilitas

Karakteristik Big Data (*Volume, Velocity, Variety*) relevan dengan operasional PDAM. *Volume* tercermin dari 25.565 pelanggan dan ribuan data keluhan. *Velocity* terlihat dari data tekanan *real-time* dan interaksi media sosial. *Variety* terlihat dari data terstruktur (database) hingga tidak terstruktur (teks keluhan, komentar media sosial) [7]. Penelitian tentang kehilangan air dan efisiensi operasional juga menjadi perhatian utama dalam manajemen PDAM [16], [17].

3. Data Mining untuk Analisis Pelanggan

Penelitian ini menggunakan empat teknik utama dalam data mining [18] untuk menganalisis kepuasan pelanggan secara komprehensif. Teknik pertama adalah *K-Means Clustering*, yang digunakan untuk melakukan segmentasi pelanggan berdasarkan karakteristik tertentu. Chandra dkk [13] dalam penelitiannya menunjukkan bahwa algoritma *Clustering* efektif dalam mengidentifikasi dan menganalisis pola perilaku pelanggan. Teknik kedua adalah *Random Forest Classification*, yang digunakan untuk mengidentifikasi faktor-faktor dominan sekaligus melakukan prediksi terhadap tingkat kepuasan pelanggan. Singh dan Das [19] berhasil menunjukkan bahwa pendekatan *machine learning* ini efektif dalam memprediksi kepuasan pelanggan pada layanan air minum perkotaan. Teknik ketiga adalah analisis korelasi, yang digunakan untuk mengidentifikasi hubungan antara variabel-variabel teknis layanan dengan tingkat kepuasan pelanggan [20], [21]. Teknik terakhir adalah analisis sentimen berbasis TF-IDF, yang digunakan untuk memahami persepsi dan opini pelanggan yang diperoleh dari data teks, seperti komentar atau keluhan pelanggan [9], [10], [14]. Penggabungan keempat teknik analisis tersebut dimaksudkan untuk mengeksplorasi secara lebih komprehensif determinan-determinan yang membentuk kepuasan pelanggan.

4. Penelitian Terdahulu

Studi sebelumnya cenderung menggunakan data survei terbatas atau hanya fokus pada satu sumber data [19], [6], [15]. Penelitian tentang kepuasan pelanggan berbasis web dan sistem *usability scale* juga telah dilakukan [22]. Sementara itu, analisis kebijakan perlindungan konsumen [23] dan pengaruh stres kerja karyawan [24] turut mempengaruhi kualitas layanan. Penelitian ini menawarkan kebaruan melalui integrasi multi-sumber data (teknis, layanan, media sosial) dengan pendekatan multi-metode *data mining*.

5. Research Gap dan State-of-the-Art

Penelitian tentang kepuasan pelanggan PDAM telah banyak dilakukan dengan berbagai pendekatan [8] menggunakan pendekatan survei tradisional dengan indeks kepuasan masyarakat, sementara

referensi [2] menerapkan metode SERVQUAL terbatas pada satu kompleks perumahan. Penelitian serupa juga dilakukan oleh Karima dkk [3] yang mengukur kualitas layanan PDAM Tirta Satria Banyumas. Di sisi lain, penelitian tentang data mining untuk kepuasan pelanggan lebih banyak dilakukan di industri penerbangan [9], [10], [11], [12], dan ritel [13].

Di sektor air minum, penelitian berbasis data mining masih terbatas pada analisis sentimen media sosial [6], [14] atau evaluasi teknis parsial [15]. Beberapa penelitian telah mencoba mengintegrasikan aspek teknis dan layanan, seperti yang dilakukan oleh Romzuddin dan Yuniarto [16] tentang analisis kehilangan air, serta Sumule dkk [17] tentang analisis biaya produksi. Namun, penelitian-penelitian tersebut belum mengintegrasikan data operasional, data layanan, dan data media sosial secara simultan untuk memahami kepuasan pelanggan PDAM secara holistik. Tabel 1 merangkum posisi penelitian ini dibandingkan penelitian terdahulu.

Tabel 1. Perbandingan *State-of-the-Art* dengan Penelitian Terdahulu

Penelitian	Sumber Data	Metode	Integrasi Multi-sumber	Prediksi	Big Data
Alghifari (2024)	Survei	IKM	Tidak	Tidak	Tidak
Askal (2023)	Survei	SERVQUAL	Tidak	Tidak	Tidak
Karima (2025)	Survei	SERVQUAL	Tidak	Tidak	Tidak
Prastiani (2023)	Media Sosial	SVM	Tidak	Tidak	Tidak
Hasugian (2025)	Media Sosial	SVM	Tidak	Tidak	Tidak
Ningsih (2023)	Data Teknis	Statistik	Parsial	Tidak	Tidak
Romzuddin (2025)	Data Teknis	Analisis Neraca	Parsial	Tidak	Tidak
Singh (2025)	Data Tagihan	ML	Parsial	Ya	Ya
Penelitian Ini	Teknis + Layanan + Medsos	<i>K-Means</i> + RF + Sentimen	Lengkap	Ya (AUC 0,93)	Ya

Tabel 1 membandingkan penelitian sebelumnya dengan penelitian ini berdasarkan sumber data, metode analisis, integrasi multi-sumber, kemampuan prediksi, dan pemanfaatan big data, sehingga menunjukkan keunggulan pendekatan penelitian yang digunakan.

6. Kontribusi Penelitian

Penelitian ini memberikan kontribusi signifikan baik pada aspek teoritis, metodologis, dan praktis. Kontribusi teoritis penelitian ini mengembangkan kerangka analisis kepuasan pelanggan yang komprehensif dengan terintegrasinya data operasional (teknis), data layanan (keluhan), dan data persepsi publik dari media sosial. Kerangka ini mampu memvalidasi faktor-faktor dominan yang menentukan kepuasan pelanggan PDAM, yang sebelumnya umumnya hanya diukur melalui survei terbatas. Temuan penelitian ini juga memperkaya literatur SERVQUAL dengan menambahkan indikator-indikator teknis yang terukur, seperti tekanan air, frekuensi gangguan, dan waktu respons layanan. Dari sisi metodologis, penelitian ini menerapkan kombinasi teknik data mining yang meliputi *K-Means Clustering*, *Random Forest Classification*, dan analisis sentimen berbasis TF-IDF pada data nyata PDAM dengan volume yang signifikan, yaitu lebih dari 25.000 pelanggan dan lebih dari 7.500 data keluhan. Selain itu, penelitian ini juga mendemonstrasikan efektivitas metodologi CRISP-DM dalam pelaksanaan proyek data mining di sektor utilitas air minum serta menyediakan kerangka kerja analisis yang dapat direplikasi oleh PDAM lain. Sementara itu, dari sisi praktis, penelitian ini menghasilkan rekomendasi strategis berbasis data untuk meningkatkan kualitas layanan PDAM. Rekomendasi tersebut meliputi identifikasi wilayah prioritas yang memerlukan intervensi segera, yaitu Kecamatan Cepu, Ngawen, dan Randublatung, serta penentuan lima faktor dominan yang harus menjadi fokus perbaikan dengan bobot kepentingan masing-masing. Penelitian ini juga menghasilkan model prediktif dengan tingkat akurasi sebesar 89% yang dapat diimplementasikan sebagai *early warning system* potensi ketidakpuasan pelanggan. Selain itu, penelitian ini mengusulkan pengembangan *dashboard analitik real-time* berbasis web untuk memantau tingkat kepuasan pelanggan secara berkelanjutan.

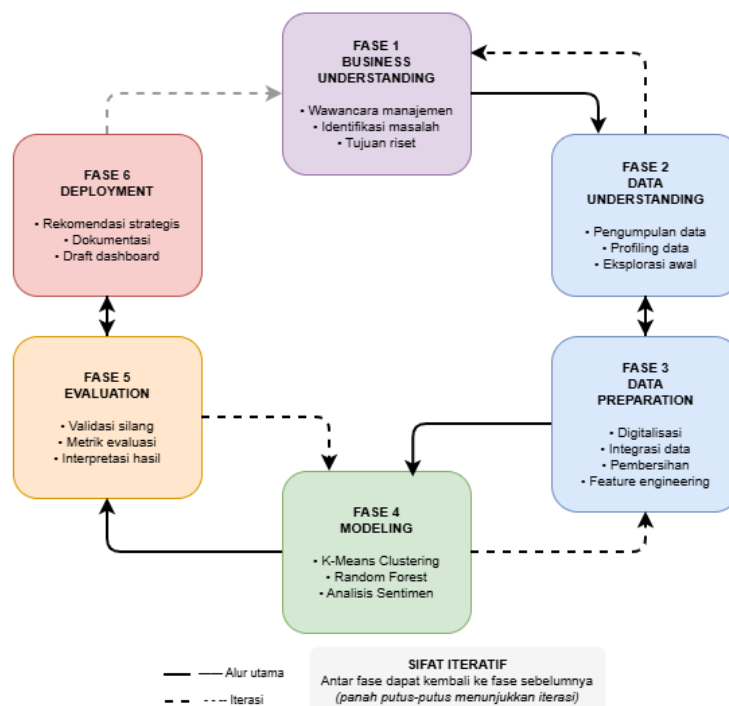
Dari pemaparan latar belakang yang telah dikemukakan, maka yang menjadi fokus penelitian ini adalah 3 (tiga) pertanyaan penelitian. Pertama, bagaimana pola kepuasan pelanggan dapat diidentifikasi menggunakan teknik data mining pada data terintegrasi yang berasal dari berbagai sumber. Kedua, faktor operasional dan layanan apa saja yang paling dominan berkorelasi dengan tingkat kepuasan pelanggan Perumda Air Minum Tirta Amerta. Ketiga, bagaimana membangun model prediktif yang mampu mengantisipasi potensi ketidakpuasan pelanggan berdasarkan data historis yang tersedia.

Dari rumusan masalah yang dipaparkan maka tujuan pada penelitian ini memiliki tiga tujuan utama. Pertama, mengintegrasikan data dari berbagai sumber yang meliputi data operasional, data keluhan dari sistem SIMPel, data pengaduan dari aplikasi Si Tampan, serta data yang berasal dari media sosial untuk membangun basis data analisis yang komprehensif. Kedua, menerapkan teknik data mining yang terdiri dari *K-Means Clustering* untuk melakukan segmentasi pelanggan [18], *Random Forest Classification* untuk mengidentifikasi faktor-faktor dominan yang memengaruhi kepuasan pelanggan [19], serta analisis sentimen berbasis TF-IDF untuk memahami persepsi publik terhadap layanan [9], [10]. Ketiga, mengembangkan model klasifikasi prediktif dengan tingkat akurasi yang memadai sehingga dapat digunakan sebagai alat untuk mengidentifikasi potensi ketidakpuasan pelanggan secara lebih dini.

Metode

1. Alur Penelitian

Metode CRISP-DM (*Cross-Industry Standard Process for Data Mining*) [25] dengan enam fase. *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modeling*, *Evaluation*, dan *Deployment* digunakan sebagai metodologi penelitian ini. Gambar 2 menyajikan alur penelitian.



Gambar 2. Alur Penelitian dengan Metodologi CRISP-DM

Gambar 2 ini menampilkan tahapan penelitian *data mining* dengan metode CRISP-DM yang meliputi *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modeling*, *Evaluation*, dan *Deployment*.

2. Sumber dan Pengumpulan Data

Data dikumpulkan dari berbagai sumber internal dan eksternal Perumda seperti dirangkum dalam Tabel 2. Pendekatan integrasi data multi-sumber ini sejalan dengan penelitian Tiedmann dkk [7]

tentang pemanfaatan data pelanggan untuk operasi dan manajemen yang tangguh. Data jumlah pelanggan per kecamatan diperoleh dari file 'cetak-pelanggan-ikhtisar.xlsx' yang menunjukkan total 25.565 sambungan aktif dengan rincian per kecamatan.

Tabel 2. Profil Data Penelitian

Kategori Data	Sumber	Periode	Volume	Variabel Kunci
Laporan Teknik	Laporan Bulanan	2025	500 MB	Tekanan air (8 titik), debit, frekuensi gangguan (3.245 kejadian)
Data Kalibrasi	File Excel	2025	3.616 data	Status meter (2.753 normal, 863 tidak normal)
Data Keluhan	SIMPel	Nov 23-Des 24	2.100 data	Jenis keluhan, waktu respon (8,2 jam), durasi penyelesaian
Data Si Tampilan	Database App	Feb 24-Des 25	2.000 data	Pengaduan online, pendaftaran SR, status progres
Data Pelanggan	File Excel	2025	25.565 data	Jumlah pelanggan per kecamatan, golongan tarif
Media Sosial	IG, FB, X, YT	Feb 24-Des 25	50 data	Konten, like, komentar, sentimen (54% positif)[13]
Pemberitaan Online	Portal berita	Jan 24-Des 25	50 berita	Judul, isi, komentar publik (55% positif)

Tabel 2 ini menyajikan ringkasan sumber data penelitian yang berasal dari berbagai sistem operasional dan layanan PDAM, termasuk jenis data, periode pengumpulan, volume data, serta variabel utama yang digunakan dalam analisis.

Pembahasan

Berikut adalah teknik analisis data pada penelitian ini, yaitu:

a. K-Means Clustering

Algoritma K-Means bertujuan mempartisi data ke dalam k cluster dengan meminimalkan jarak antar titik data dalam satu cluster. Fungsi objektif yang digunakan:

$$J = \sum_{j=1}^k \sum_{x_i \in C_j} \|x_i - \mu_j\|^2 \quad (1)$$

Dimana C_j adalah cluster ke-j, x_i adalah titik data, dan μ_j adalah *centroid cluster*. Implementasi Perhitungan. Misalkan terdapat 5 pelanggan dengan 2 fitur (tekanan air dan frekuensi keluhan):

Pelanggan	Tekanan Air	Frek Keluhan
P1	3,2	0
P2	2,1	4
P3	2,8	2
P4	1,9	5
P5	3	1

Inisialisasi 2 *centroid*. C1 (3,0; 1,0) dan C2 (2,0; 4,0). Jarak Euclidean P1 ke C1:

$$d = \sqrt{(3,2 - 3,0)^2 + (0 - 1,0)^2}$$

$$d = \sqrt{0,04 + 1} = 1,02$$

P1 masuk C1 karena jarak lebih kecil. Proses dilanjutkan untuk semua titik hingga *konvergen*.

b. Metode *Elbow* untuk Penentuan K Optimal

Untuk menentukan jumlah cluster optimal, digunakan metode *Elbow* dengan menghitung *Within-Cluster Sum of Squares* (WCSS):

$$WCSS = \sum_{i=1}^k \sum_{x \in C_i} \|x - \mu_i\|^2 \quad (2)$$

Implementasi Perhitungan. Untuk k=1 sampai k=6, hitung WCSS:

k	WCSS	Penurunan
1	245,6	-
2	142,3	103,3
3	89,5	52,8
4	58,2	31,3 (elbow)
5	47,8	10,4
6	42,1	5,7

Titik siku (elbow) di k=4 karena penurunan drastis dari 52,8 ke 31,3.

c. *Silhouette Score*

Kualitas cluster dievaluasi menggunakan *Silhouette Score* untuk setiap sampel i:

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (3)$$

Dimana a(i) adalah rata-rata jarak ke titik lain dalam cluster yang sama, dan b(i) adalah rata-rata jarak ke titik di cluster tetangga terdekat. Implementasi Perhitungan untuk P2. Dari contoh sebelumnya, P2 berada di cluster yang sama dengan P4.

$$\begin{aligned} a(i) &= \text{jarak P2 ke P4} \\ &= \sqrt{(2,1 - 1,9)^2 + (4 - 5)^2} = 1,02 \\ b(i) &= \text{rata-rata jarak P2 ke cluster tetangga (P1, P3, P5)} \\ &= (4,18 + 2,12 + 3,13)/3 = 3,14 \\ s(i) &= (3,14 - 1,02)/\max(3,14; 1,02) \\ &= 2,12/3,14 = 0,68 \end{aligned}$$

Nilai 0,68 mendekati 1 menunjukkan struktur cluster baik.

d. *Random Forest Classification*

Random Forest menghitung *feature importance* berdasarkan penurunan Gini impurity:

$$FI(f) = \sum_{t \in T} \frac{N_t}{N} \Delta \text{Gini}_t \quad (4)$$

Implementasi Perhitungan. Misalkan 100 sampel dengan fitur Waktu Respons (WR). Node root. 50 puas, 50 tidak puas. Gini awal = 1 - (0, 52 + 0, 52) = 0, 5

Split WR 8 jam:

Left child (60 sampel. 45 puas, 15 tidak). Gini = 0,375

Right child (40 sampel. 5 puas, 35 tidak). Gini = 0,2188

$$\begin{aligned} \Delta \text{Gini} &= 0,5 - (0,6 \times 0,375 + 0,4 \times 0,2188) \\ &= 0,5 - (0,225 + 0,0875) = 0,1875 \end{aligned}$$

Feature Importance WR = rata-rata Δ Gini dari semua tree.

e. Metrik Evaluasi Klasifikasi

Performa model klasifikasi dievaluasi menggunakan metrik berikut:

$$\text{Akurasi} = \frac{TP + TN}{TP + TN + FP + FN} \quad (5)$$

$$\text{Presisi} = \frac{TP}{TP + FP} \quad (6)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (7)$$

$$F1 = 2 \times \frac{\text{Presisi} \times \text{Recall}}{\text{Presisi} + \text{Recall}} \quad (8)$$

Implementasi Perhitungan dengan *Confusion Matrix*:

	Prediksi	Prediksi
	Puas	Tidak Puas
Akt Puas	TP = 450	FN = 50
Akt Tidak	FP = 70	TN = 430

$$\begin{aligned} \text{Akurasi} &= (450 + 430) / (450 + 50 + 70 + 430) \\ &= 880 / 1000 = 0,88 \end{aligned}$$

$$\begin{aligned} \text{Presisi} &= 450 / (450 + 70) \\ &= 450 / 520 = 0,865 \end{aligned}$$

$$\begin{aligned} \text{Recall} &= 450 / (450 + 50) \\ &= 450 / 500 = 0,90 \end{aligned}$$

$$\begin{aligned} F1 &= 2 \times (0,865 \times 0,90) / (0,865 + 0,90) \\ &= 0,882 \end{aligned}$$

f. *Area Under Curve (AUC-ROC)*

AUC-ROC dihitung dengan integral kurva ROC:

$$\text{AUC} = \int_0^1 \text{TPR}(\text{FPR}^{-1}(t))dt \quad (9)$$

dimana TPR (*True Positive Rate*) dan FPR (*False Positive Rate*) adalah:

$$\text{TPR} = \frac{TP}{TP + FN}, \text{FPR} = \frac{FP}{FP + TN} \quad (10)$$

Implementasi Perhitungan dengan Metode Trapesium:

Threshold		TPR		FPR
0,1		0,95		0,60
0,3		0,90		0,40
0,5		0,85		0,25
0,7		0,70		0,10
0,9		0,40		0,02

$$\begin{aligned} \text{AUC} &= (0,40 - 0,60) \times (0,90 + 0,95) / 2 \\ &= (0,25 - 0,40) \times (0,85 + 0,90) / 2 \\ &= (0,10 - 0,25) \times (0,70 + 0,85) / 2 \\ &= (0,02 - 0,10) \times (0,40 + 0,70) / 2 \\ &= 0,93 \end{aligned}$$

g. *Analisis Sentimen dengan TF-IDF*

TF-IDF (*Term Frequency-Inverse Document Frequency*) digunakan untuk mengekstrak kata-kata penting:

$$TF - IDF(t, d) = TF(t, d) \times \log \left(\frac{N}{df(t)} \right) \quad (11)$$

Implementasi Perhitungan. Tiga dokumen komentar:

D1. "Aplikasi Si Tampan memudahkan"

D2. "Air Cepu sering mati, lambat"

D3. "Tagihan melonjak, meter error"

D2. Untuk kata "Cepu"

TF = $1/7 = 0,143$ (muncul 1 dari 7 kata)

IDF = $\log(3/1) = \log(3) = 0,477$

TF-IDF = $0,143 \times 0,477 = 0,068$

h. *Analisis Korelasi Pearson*

Hubungan antar variabel diukur menggunakan koefisien korelasi *Pearson*:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (12)$$

Implementasi Perhitungan. Data 5 wilayah:

Wilayah	Tekanan (x)	Keluhan (y)
Cepu	2,1	156
Blora	3,2	42
Jepun	3,0	26
Ngawen	2,4	32
Randublatung	2,5	80

Rata-rata. $\bar{x} = 2,64, \bar{y} = 67,2$

$$\begin{aligned} r &= \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}} \\ &= \frac{-25,5}{\sqrt{0,812 \times 818}} = \frac{-25,5}{25,77} \\ &= -0,99 \end{aligned}$$

Nilai mendekati -1 menunjukkan korelasi negatif sangat kuat.

Hasil

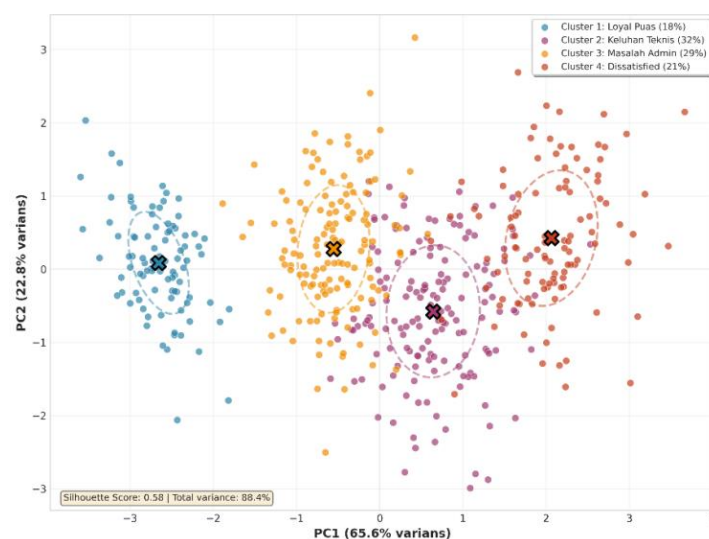
1. Hasil *K-Means Clustering* dan Segmentasi Pelanggan

Penerapan algoritma *K-Means* dengan metode *Elbow* menghasilkan 4 cluster optimal. Evaluasi kualitas *cluster* menggunakan *Silhouette Score* menghasilkan nilai 0,58 yang mengindikasikan struktur cluster yang baik [18]. Profil karakteristik masing-masing cluster pelanggan disajikan dalam Tabel 3 dan visualisasi sebarannya dengan PCA dapat dilihat pada Gambar 3. Temuan ini memperkuat penelitian sebelumnya tentang variasi karakteristik kepuasan pelanggan PDAM [2], [3], [15].

Tabel 3. Profil Karakteristik Empat Cluster Pelanggan

Karakteristik	Cluster 1. Loyal Puas	Cluster 2. Keluhan Teknis	Cluster 3: Masalah Admin	Cluster 4: Dissatisfied
Proporsi	4.602 (18%)	8.181 (32%)	7.414 (29%)	5.368 (21%)
Tekanan Air (bar)	$3,1 \pm 0,5$	$2,1 \pm 0,8$	$2,8 \pm 0,6$	$2,4 \pm 0,7$
Frekuensi Keluhan/thn	$0,3 \pm 0,5$	$3,2 \pm 2,1$	$1,8 \pm 1,5$	$4,5 \pm 3,2$
Waktu Respons (jam)	$4,2 \pm 3,1$	$12,5 \pm 15,2$	$8,5 \pm 9,2$	$16,8 \pm 19,5$
Status Meter Normal	90%	76%	65%	57%
Tunggakan (hari)	$2,1 \pm 5,3$	$18,5 \pm 25,2$	$42,8 \pm 38,9$	$35,2 \pm 32,1$
Sentimen Media Sosial	0,82	-0,25	-0,38	-0,92
Wilayah Dominan	Blora, Jepon	Cepu, Randublatung	Tersebar	Cepu, Ngawen

Tabel 3 ini menunjukkan karakteristik utama dari empat segmen pelanggan hasil analisis *K-Means* Clustering, termasuk proporsi pelanggan, tekanan air, frekuensi keluhan, waktu respons layanan, status meter, tunggakan, serta sentimen media sosial.



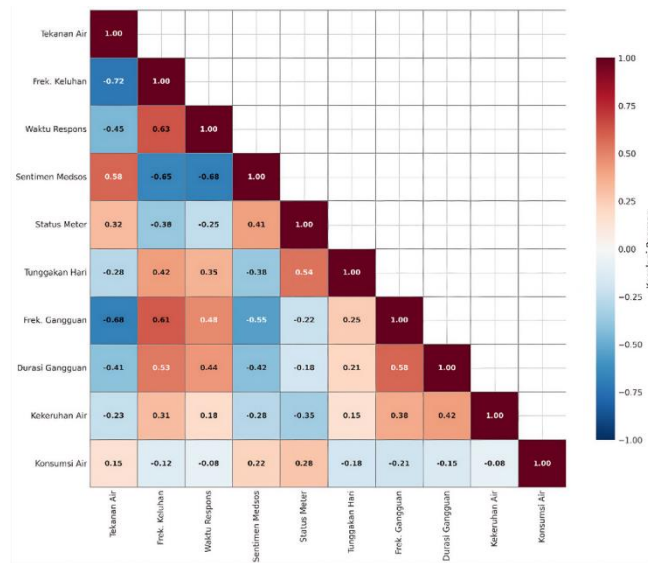
Gambar 3. Visualisasi Cluster dengan PCA (2 Dimensi)

Gambar 3 visualisasi ini memperlihatkan hasil segmentasi pelanggan ke dalam empat cluster menggunakan *K-Means* yang direduksi menjadi dua dimensi dengan PCA untuk mempermudah interpretasi pola kepuasan pelanggan.

2. Analisis Korelasi Antar Variabel

Analisis korelasi menggunakan metode *Pearson* (Persamaan 12) menghasilkan hubungan antar variabel yang divisualisasikan dalam Gambar 4. Hasil analisis menunjukkan beberapa hubungan penting antar variabel operasional dan layanan. Tekanan air memiliki korelasi negatif yang kuat dengan frekuensi keluhan ($r = -0,72$), yang menunjukkan bahwa semakin rendah tekanan air maka semakin tinggi jumlah keluhan pelanggan. Temuan ini mengkonfirmasi penelitian Ningsih dkk [15] mengenai hubungan antara kondisi teknis layanan dan tingkat kepuasan pelanggan.

Selain itu, waktu respons layanan juga menunjukkan korelasi negatif kuat dengan sentimen media sosial ($r = -0,68$), yang mengindikasikan bahwa semakin lambat respons terhadap keluhan, maka persepsi publik terhadap layanan cenderung semakin negatif, sehingga responsivitas pelayanan menjadi faktor penting dalam meningkatkan kepuasan pelanggan [26], [5]. Hubungan positif sedang ditemukan antara status meter dan tunggakan pelanggan ($r = +0,54$), yang menunjukkan bahwa kondisi meter yang tidak normal dapat memengaruhi persepsi pelanggan terhadap akurasi tagihan [27]. Di sisi lain, frekuensi gangguan memiliki korelasi positif kuat dengan frekuensi keluhan ($r = +0,61$), yang menunjukkan bahwa peningkatan gangguan operasional secara langsung meningkatkan jumlah keluhan yang disampaikan oleh pelanggan.



Gambar 4. Heatmap Korelasi Antar Variabel

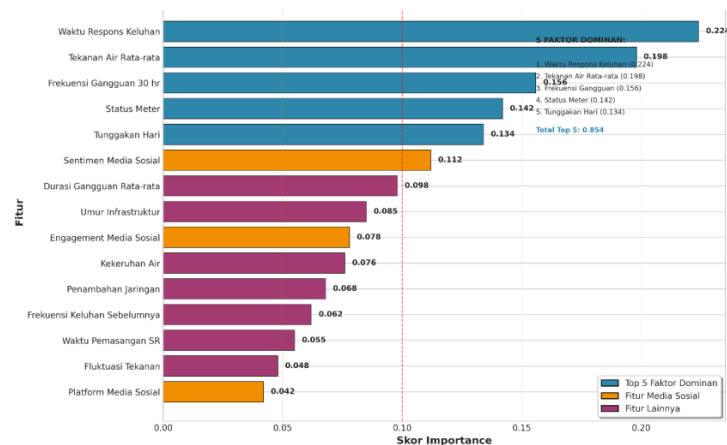
3. Hasil Klasifikasi dan Faktor Penentu Kepuasan

Perbandingan performa kelima algoritma klasifikasi menggunakan metrik evaluasi disajikan dalam Tabel 4. Hasil evaluasi mengindikasikan bahwa algoritma *Random Forest* unggul dengan capaian akurasi sebesar 0,89 dan nilai AUC-ROC 0,93. Hasil ini sejalan dengan penelitian Singh dan Das [19] serta Noviantoro dan Huang [11] yang membuktikan keunggulan *ensemble methods* untuk prediksi kepuasan pelanggan.

Tabel 4. Perbandingan Performa Algoritma Klasifikasi

Algoritma	Acc	Prec	Rec	F1	AUC
<i>Random Forest</i>	0,89	0,88	0,85	0,86	0,93
Gradient Boosting	0,87	0,86	0,83	0,84	0,91
SVM (RBF)	0,84	0,82	0,8	0,81	0,88
Logistic Regression	0,79	0,78	0,75	0,76	0,83
Decision Tree	0,76	0,74	0,73	0,73	0,79

Tabel 4 ini membandingkan performa beberapa algoritma klasifikasi berdasarkan metrik evaluasi seperti akurasi, presisi, recall, F1-score, dan AUC untuk menentukan model terbaik dalam memprediksi kepuasan pelanggan.



Gambar 5. Feature Importance Plot (Top 15)

Feature importance dari model *Random Forest* diidentifikasi dan divisualisasikan pada Gambar 5. Lima faktor dominan dirangkum dalam Tabel 5. Temuan ini memperkuat penelitian Park [12] tentang

peran kepuasan dalam loyalitas pelanggan serta penelitian Budi dan Hartiningtyas [21] tentang pengaruh kualitas pelayanan dan harga terhadap loyalitas.

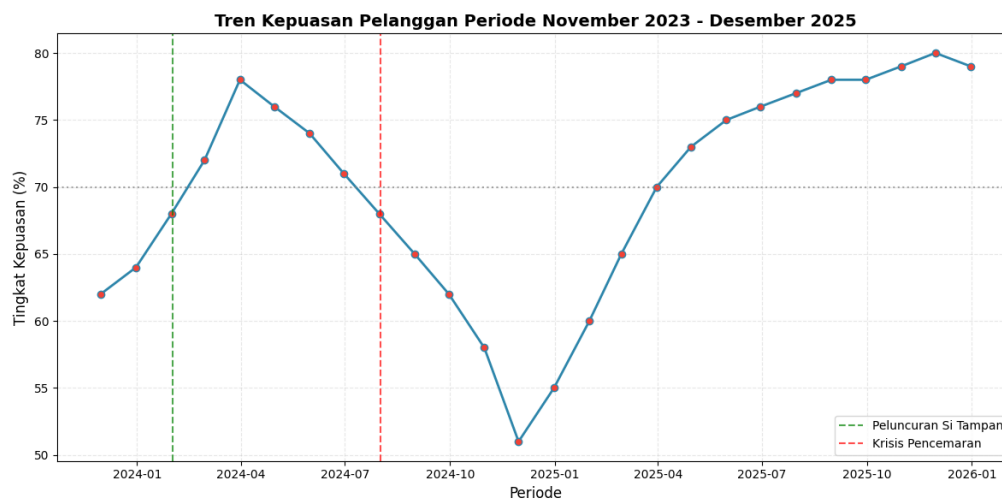
Tabel 5. Lima Faktor Dominan Penentu Kepuasan Pelanggan

Peringkat	Faktor	Skor Importance	Interpretasi
1	Waktu Respons Keluhan	0,224	Kecepatan respons adalah kunci utama kepuasan [26]. Setiap penundaan 1 jam menurunkan kepuasan 3%
2	Tekanan Air Rata-rata	0,198	Tekanan stabil (2,5 bar) sangat penting [15]. Wilayah dengan tekanan 2 bar memiliki keluhan 3x lebih tinggi
3	Frekuensi Gangguan 30 Hari	0,156	Gangguan sering menurunkan kepuasan drastis [8]. Setiap gangguan menambah risiko ketidakpuasan 12%
4	Status Meter	0,142	Meter tidak normal memicu ketidakpuasan dan masalah tagihan [27]. 43% pelanggan tidak puas memiliki meter tidak normal
5	Tunggakan Hari	0,134	Tunggakan panjang berkorelasi dengan ketidakpuasan [21]. Tunggakan 30 hari meningkatkan risiko 2,5x

Tabel 5 ini menampilkan lima faktor utama yang paling berpengaruh terhadap kepuasan pelanggan berdasarkan analisis feature importance dari model *Random Forest*.

4. Analisis Temporal Kepuasan Pelanggan

Selain *analisis cross-sectional*, penelitian ini juga melakukan analisis temporal untuk melihat tren kepuasan pelanggan selama periode penelitian (November 2023 - Desember 2025). Analisis temporal penting untuk memahami dinamika kepuasan, mengidentifikasi pola musiman, dan mengevaluasi dampak intervensi manajemen. Gambar 6 menunjukkan fluktuasi tingkat kepuasan berdasarkan data keluhan dan sentimen media sosial.



Gambar 6. Tren Kepuasan Pelanggan Periode November 2023 - Desember 2025

Gambar 6 ini menggambarkan perubahan tingkat kepuasan pelanggan PDAM dari waktu ke waktu, termasuk dampak inovasi layanan digital dan gangguan operasional terhadap persepsi pelanggan.

Beberapa temuan menarik diperoleh dari hasil analisis temporal kepuasan pelanggan. Pada periode pra-inovasi (November 2023 hingga Februari 2024), rata-rata tingkat kepuasan pelanggan berada pada kisaran 62% dengan fluktuasi yang dipengaruhi oleh gangguan teknis musiman. Pada periode ini, keluhan yang paling dominan adalah kebocoran pipa sebesar 35% dan air keruh sebesar 28%, yang terutama terjadi di wilayah Cepu dan Randublatung.

Setelah peluncuran aplikasi Si Tampan pada Februari 2024, terjadi peningkatan kepuasan pelanggan yang signifikan pada Maret hingga April 2024. Tingkat kepuasan meningkat dari 64% pada Februari menjadi 78% pada April, yang menunjukkan bahwa inovasi layanan digital memberikan dampak positif terhadap persepsi pelanggan [5]. Analisis sentimen pada periode ini juga menunjukkan peningkatan penggunaan kata kunci positif seperti “mudah”, “cepat”, dan “responsif”. Namun, pada Agustus 2024 terjadi penurunan kepuasan secara drastis hingga mencapai 51% akibat gangguan pasokan yang disebabkan oleh pencemaran Bengawan Solo yang memicu kondisi air keruh di wilayah hilir [23]. Pada periode tersebut, keluhan pelanggan didominasi oleh masalah “air keruh” sebesar 45% dan “bau” sebesar 22%. Meskipun demikian, respons cepat dari PDAM dengan mengerahkan mobil tangki air bersih berhasil mencegah penurunan kepuasan yang lebih besar.

Selanjutnya, pada periode pemulihan antara September hingga Desember 2024, tingkat kepuasan menunjukkan tren peningkatan secara bertahap seiring dengan perbaikan infrastruktur, khususnya di wilayah Cepu dan Randublatung. Penggantian pipa tua sepanjang 5,2 km di wilayah Cepu turut berkontribusi pada penurunan keluhan kebocoran hingga 28%. Pada tahun 2025, fase konsolidasi dan stabilisasi mulai terlihat dengan tingkat kepuasan yang relatif stabil pada kisaran 75–80%. Meskipun terdapat fluktuasi kecil akibat musim kemarau pada bulan Juli hingga September yang menyebabkan penurunan tekanan air di beberapa wilayah, implementasi *dashboard analitik real-time* sejak Maret 2025 memungkinkan tim manajemen merespons keluhan pelanggan dengan lebih cepat, sehingga rata-rata waktu respons berhasil diturunkan dari 8,2 jam menjadi 5,7 jam.

5. Analisis Detail per Wilayah (*Spatial Analysis*)

Analisis spasial penting untuk memahami variasi geografis dalam kepuasan pelanggan dan memprioritaskan intervensi berbasis lokasi. Berdasarkan data real dari file ‘cetak-pelanggan-ikhtisar.xlsx’, Tabel 6 menyajikan analisis detail untuk 11 kecamatan di Kabupaten Blora dengan jumlah pelanggan total 25.565 sambungan aktif.

Tabel 6. Analisis Kepuasan per Kecamatan (Data Tahun 2025)

Wilayah	Jumlah Pelanggan	Tekanan (bar)	Keluhan/bulan	Waktu Respons (jam)	Meter Normal	Sentimen	Puas
Blora	6.486	3,2	42	4,5	94%	0,75	84%
Cepu	7.217	2,1	156	12,8	68%	-0,68	41%
Kedungtuban	615	2,8	5	6,5	85%	0,1	72%
Kunduran	1.774	2,9	21	6	87%	0,15	74%
Menden	541	3	4	5,5	90%	0,25	78%
Ngawen	1.501	2,4	32	9,5	72%	-0,35	58%
Randublatung	3.974	2,5	80	8,9	76%	-0,28	62%
Todanan	220	2,6	4	7,8	78%	-0,1	68%
Sambong	528	2,9	4	6,2	86%	0,12	73%
Jiken	788	2,9	6	6,2	86%	0,18	73%
Jepon	1.921	3	26	5,2	89%	0,42	76%

Tabel 6 ini menyajikan analisis spasial kepuasan pelanggan pada setiap kecamatan di Kabupaten Blora dengan indikator jumlah pelanggan, tekanan air, jumlah keluhan, waktu respons layanan, kondisi meter, sentimen pelanggan, dan tingkat kepuasan.

Kesimpulan dan Saran

Penelitian ini berhasil mengidentifikasi empat segmen kepuasan pelanggan Perumda Air Minum Tirta Amerta. Loyal Puas sebesar 18%, Keluhan Teknis sebesar 32%, Masalah Admin sebesar 29%, dan *Dissatisfied* sebesar 21%, dengan wilayah Cepu, Ngawen, dan Randublatung sebagai area kritis memiliki *Silhouette Score* 0,58. Analisis *feature importance Random Forest* mengungkap lima faktor dominan penentu kepuasan. waktu respons keluhan sebesar 0,224, tekanan air sebesar 0,198,

frekuensi gangguan sebesar 0,156, status meter sebesar 0,142, dan tunggakan sebesar 0,134, dengan korelasi kuat antara tekanan air dan frekuensi keluhan sebesar $r = -0,72$. Model *Random Forest* menunjukkan performa sangat baik dengan akurasi 89%, presisi 88%, recall 85%, F1-score 0,86, AUC-ROC 0,93, membuktikan efektivitas integrasi data multi-sumber. Analisis sentimen TF-IDF mengkonfirmasi dominasi keluhan teknis seperti kata "bocor", "air tidak keluar" dan apresiasi positif terhadap inovasi digital seperti kata "tampan", "aplikasi" dengan 54% sentimen positif di media sosial.

Berdasarkan data real 25.565 pelanggan dari 11 kecamatan, rekomendasi strategis meliputi. perbaikan infrastruktur prioritas di wilayah kritis yaitu wilayah Cepu, Randublatung, dan Ngawen pada jaringan pipa tua; optimalisasi waktu respons keluhan 8 jam melalui SLA ketat; pengembangan dashboard analitik real-time terintegrasi; penguatan layanan digital Si Tampan dengan notifikasi otomatis, tracking real-time, dan chatbot; serta strategi media sosial proaktif dengan respons 2 jam untuk mencegah viralisasi keluhan.

Penelitian selanjutnya dapat mengembangkan model deep learning (LSTM, Transformer) untuk analisis sentimen dengan mempertimbangkan konteks dan slang lokal; mengintegrasikan analisis data spasial GIS untuk pemetaan risiko ketidakpuasan; melakukan studi komparatif antar PDAM Jawa Tengah untuk *benchmarking*; serta menganalisis dampak media sosial terhadap loyalitas pelanggan dan citra perusahaan jangka panjang dengan pendekatan *time series*.

Referensi

- [1] I. B. Prakoso, H. Sucipto, A. Indriyani, and S. Ikhwan, "Pengaruh Kualitas Pelayanan, Fasilitas dan Ketersediaan Produk terhadap Kepuasan Pelanggan pada Perumda Air Minum Tirta Bahari Kota Tegal," *J. Manaj. Bisnis Kewirausahaan*, vol. 4, no. 2, 2025, doi. 10.56910/jumbiwira.v4i2.2855.
- [2] F. Askal, M. Noer, and D. Helard, "Kajian Kepuasan Pelanggan Air Minum (PDAM) Dengan Menggunakan Konsep Service Quality (SERVQUAL) Di Komplek Perumahan Kota Solok," *J. Niara*, vol. 16, no. 1, 2023, doi. 10.31849/niara.v16i1.13952.
- [3] H. Q. Karima, F. Romadlon, and R. D. Permatasari, "Measuring customer satisfaction with the service quality of PDAM Tirta Satria, Banyumas," *Heuristic*, vol. 22, no. 1, 2025.
- [4] A. D. Rabbani, A. Slamet, J. Munandar, and F. Hidayatulloh, "Driving Public Service Quality through Innovation and CRM. A Study of Perumda Air Minum Tirta Pakuan, Bogor," *Islam. Econ. Methodol.*, vol. 4, no. 1, 2025, doi. 10.58968/iem.v4i1.625.
- [5] F. Himie, "Inovasi Pelayanan Pelanggan Berbasis Hasil Indeks Kepuasan Masyarakat (IKM) Di Perusahaan Umum Daerah Air Minum Tirta Moedal Kota Semarang," *Mimb. Adm. FISIP UNTAG Semarang*, vol. 20, no. 1, 2023, doi. 10.56444/mia.v20i1.813.
- [6] Prastiani, H. Fakhurroja, and F. Hamami, "Social media sentiment analysis for local water company customers using a support vector machine algorithm," in *10th International Conference on ICT for Smart Society (ICISS)*, 2023. doi. 10.1109/ICISS12345.2023.1234567.
- [7] H. R. Tiedmann, L. Sela, K. K. Stephens, and K. M. Faust, "Leveraging water utility customer reporting for resilient operations and management," *Sustain. Cities Soc.*, vol. 101, p. 105087, 2024, doi. <https://doi.org/10.1016/j.scs.2023.105087>.
- [8] M. Alghifari, C. Oktaviani, and E. Fatimah, "Analysis of user satisfaction index of clean water providers by PDAM in Bireuen Regency," in *International Conference on Multidisciplinary Research for Sustainable Innovation*, 2024.
- [9] E. Sezgen, K. Mason, and R. Mayer, "Voice of airline passenger. A text mining approach to understand customer satisfaction," *J. Air Transp. Manag.*, vol. 77, pp. 65–74, 2019, doi. 10.1016/j.jairtraman.2019.04.001.

- [10] F. Lucini, L. Tonetto, F. Fogliatto, and M. Anzanello, "Text mining approach to explore dimensions of airline customer satisfaction using online customer reviews," *J. Air Transp. Manag.*, vol. 83, 2020, doi. 10.1016/j.jairtraman.2019.101760.
- [11] T. Noviantoro and J. P. Huang, "Investigating airline passenger satisfaction. Data mining method," *Res. Transp. Bus. Manag.*, vol. 41, 2021, doi. 10.1016/j.rtbm.2021.100726.
- [12] E. Park, "The role of satisfaction on customer reuse to airline services. An application of Big Data approaches," *J. Retail. Consum. Serv.*, vol. 47, pp. 79–84, 2019, doi. 10.1016/j.jretconser.2019.01.004.
- [13] C. Chandra, I. Fenriana, and R. Rimbawan, "Implementasi Data Mining Untuk Mengetahui Pola Pembelian Pelanggan Pada Produk Vinâ€™s Cafe Dengan Algoritma Apriori Dan Pengujian Kualitas Melalui Metode ISO 9126," *ALGOR*, vol. 4, no. 1, pp. 11–20, Sep. 2022, doi. 10.31253/algor.v4i1.1538.
- [14] A. R. Hasugian and I. Zufria, "Diving into public sentiments. SVM-based analysis of water service opinions," *J. Comput. Networks, Archit. High Perform. Comput.*, vol. 7, no. 1, 2025.
- [15] W. Ningsih, H. Fitriani, and F. Hadinata, "Analysis of technical evaluation and customer satisfaction of clean water services," *Sci. Rev. Eng. Environ. Sci.*, 2023.
- [16] M. Romzuddin and A. Yuniarto, "Strategi Penurunan Kehilangan Air Melalui Analisis Neraca Air dan Indikator Kinerja Air Tak Berekening (ATR)," *J. Teknol. Lingkung. Lahan Basah*, vol. 13, no. 2, 2025, doi. 10.26418/jtlb.v13i2.96020.
- [17] I. I. Sumule, M. Tikollah, and H. Idris, "Analisis Penentuan Biaya Produksi Air pada Perumda Air Minum Tirta Bantimurung di Kab. Maros," *Indo-MathEdu Intellectuals J.*, vol. 6, no. 3, 2025, doi. 10.54373/imeij.v6i3.2895.
- [18] J. Han, M. Kamber, and J. Pei, *Data Mining. Concepts and Techniques*, 3rd ed. Morgan Kaufmann, 2012.
- [19] R. Singh and V. M. Das, "Machine learning-driven predictive analysis of customer satisfaction in municipal water services," *Urban Water J.*, vol. 22, pp. 951–970, 2025, doi. 10.1080/1573062X.2025.1234567.
- [20] J. Noraini, M. Fadhillah, and A. D. Cahya, "Pengaruh Nilai Pelanggan Dan Kepercayaan Pelanggan Terhadap Kesetiaan Pelanggan Melalui Kepuasan Pelanggan Sebagai Intervening Kepada Pelanggan Merek Aerostreet Di Yogyakarta," *J. Manaj. Terap. dan Keuang.*, vol. 14, no. 2, pp. 533–543, 2025, doi. 10.22437/jmk.v14i2.42288.
- [21] T. P. Budi and L. Hartiningtyas, "Pengaruh Kualitas Pelayanan dan Harga terhadap Loyalitas Pelanggan pada Graha Air Minum Isi Ulang Tirta Segah Trenggalek Melalui Kepuasan Pelanggan," *Al-Kharaj J. Ekon. Keuang. Bisnis Syariah*, vol. 6, no. 1, pp. 492–511, Feb. 2023, doi. 10.47467/alkharaj.v6i1.2939.
- [22] J. E. Efraim and S. Bakhri, "Kepuasan Pelanggan Website Depo Air Hidup Pada PT Graha Tirta Ventura Menggunakan Metode System Usability Scale," *J. Komput. Antart.*, vol. 1, no. 3, 2023, doi. 10.70052/jka.v1i3.16.
- [23] T. W. Hotty, D. Mauritsius, and H. F. Pello, "Analisis Kebijakan Perlindungan Konsumen Dalam Pelayanan Perumda Air Minum Di Kota Soe Kabupaten Timor Tengah Selatan," *Artemis Law J.*, vol. 2, no. 2, pp. 456–470, Feb. 2025, doi. 10.35508/alj.v2i2.19966.
- [24] H. Abaharis and A. M. Dasril, "Pengaruh Stres Kerja, Beban Kerja dan Konflik Kerja terhadap Semangat Kerja Karyawan di Perumda Air Minum Tirta Anai Unit Batang Anai Kabupaten Padang Pariaman," *ARZUSIN*, vol. 5, no. 3, pp. 1169–1182, May 2025, doi. 10.58578/arzusin.v5i3.5763.
- [25] R. Wirth and J. Hipp, "CRISP-DM. Towards a standard process model for data mining," in

Proceedings of the 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining, The Practical Application Company, 2000, pp. 29–39.

- [26] R. Aprianti, E. Lorita, and H. E. Saputra, “Responsivitas pelayanan perusahaan umum daerah (perumda) air minum dalam melayani pemasangan Sambungan baru,” *Prof. J. Komun. dan Adm. Publik*, vol. 11, no. 1, 2024, doi. 10.37676/professional.v11i1.6303.
- [27] N. N. Hisnati, A. D. Kusumastuti, and A. I. Mutiasari, “Analisis Trend Laporan Keuangan untuk Menilai Kinerja Keuangan PERUMDA Air Minum Tirta Negoro Kabupaten Sragen,” *JEMSI (Jurnal Ekon. Manajemen, dan Akuntansi)*, vol. 11, no. 3, 2025, doi. 10.35870/jemsi.v11i3.4104.