Department of Digital Business
**Journal of Artificial Intelligence and Digital Business (RIGGS)**
Homepage: https://journal.ilmudata.co.id/index.php/RIGGS
Vol. 4 No. 3 (2025) pp: 4831-4838
P-ISSN: 2963-9298, e-ISSN: 2963-914X

# Perbandingan Kinerja Algoritma K-Nearest Neighbor dan Naive Bayes untuk Klasifikasi Loyalitas Pelanggan (Studi Kasus: CV Cahaya Alam Indah)

Abdan Syakur Ramadhan[1], Lena Magdalena[2], Mesi Febima[3]
[1,2,3]Prodi Sistem Informasi, Fakultas Teknologi Informasi, Universitas Catur Insan Cendikia
[1]abdan.ramadhan.si.21@cic.ac.id, [2]lena.magdalena@cic.ac.id, [3]mesi.febima@cic.ac.id

***Abstrak***

*CV Cahaya Alam Indah menghadapi tantangan dalam merancang strategi pemasaran yang efektif karena tidak adanya sistem klasifikasi pelanggan yang terstruktur. Penelitian ini bertujuan untuk membangun model klasifikasi dengan membandingkan performa algoritma K-Nearest Neighbors (KNN) dan Naive Bayes untuk mengkategorikan pelanggan menjadi Loyal, Cenderung Loyal, dan Tidak Loyal. Penelitian ini menggunakan 220 data riwayat penjualan pelanggan yang melalui tahap pra-pemrosesan menggunakan normalisasi Min-Max dan dibagi menjadi 70% data latih dan 30% data uji. Kinerja model dievaluasi berdasarkan akurasi, presisi, recall, dan F1-Score. Hasil penelitian menunjukkan bahwa algoritma KNN mencapai akurasi lebih tinggi sebesar 83.3%, mengungguli algoritma Naive Bayes yang memperoleh 80.3%. Nilai F1-Score KNN juga secara konsisten lebih superior di semua kelas. Dengan demikian, model KNN direkomendasikan sebagai solusi yang lebih efektif untuk klasifikasi loyalitas pelanggan pada studi kasus ini.*

*Kata kunci: Klasifikasi, Loyalitas Pelanggan, K-Nearest Neighbors, Machine Learning, Naive Bayes.*

## 1. Latar Belakang

Di era digital, pemanfaatan data untuk klasifikasi pelanggan menjadi kunci bagi perusahaan dalam merancang strategi pemasaran yang efektif dan membangun loyalitas jangka panjang [1]. Hal ini sangat relevan bagi industri air minum galon isi ulang (DAMIU), dimana pemahaman mendalam terhadap karakteristik pelanggan yang umumnya bersifat domestik, sensitif terhadap harga, dan mengutamakan kualitas layanan menjadi faktor vital untuk menjaga daya saing [2]. Pentingnya sektor ini semakin meningkat akibat tantangan ketersediaan air bersih yang dipicu oleh pencemaran lingkungan [3], [4] dan dampak perubahan iklim [5], yang menjadikan air bersih sebagai komoditas vital [6] dan menciptakan ketidakseimbangan antara ketersediaan dengan permintaan masyarakat [7].

CV Cahaya Alam Indah, produsen air minum 'Cairo Pure Water', menghadapi tantangan signifikan dalam hal ini. Permasalahan utamanya adalah belum adanya sistem klasifikasi pelanggan yang terstruktur untuk mengkategorikan mereka sebagai Loyal, Cenderung Loyal, atau Tidak Loyal. Akar masalah ini terletak pada proses pencatatan transaksi yang masih dilakukan secara manual pada buku penjualan, sehingga data mentah yang terkumpul sulit dianalisis dan rawan kesalahan. Dampak negatifnya telah dirasakan melalui keluhan dari pelanggan setia yang merasa kontribusi mereka tidak diapresiasi. Kegagalan dalam memanfaatkan data riwayat pembelian [8], frekuensi transaksi [9], dan preferensi merek produk [10] ini secara langsung menghambat efektivitas pemasaran dan retensi pelanggan.

Untuk mengatasi permasalahan tersebut, penelitian ini mengusulkan penerapan machine learning sebagai solusi. Pendekatan ini relevan dengan penelitian sebelumnya yang telah menunjukkan keberhasilan algoritma seperti K-Nearest Neighbors (KNN) dalam analisis kepuasan pelanggan dengan akurasi tinggi [11], serta perbandingan algoritma lain dalam konteks serupa [12]. Penelitian ini akan secara spesifik membandingkan performa dua metode: K-Nearest Neighbors (KNN), yang mengklasifikasikan data berdasarkan kedekatan jarak dengan data historis, dan Naive Bayes, yang menggunakan pendekatan probabilitas statistik untuk menentukan kelas yang paling sesuai [13], [14], [15].

penelitian ini bertujuan untuk merancang dan membangun sebuah sistem yang mampu mengklasifikasikan pelanggan di CV Cahaya Alam Indah dengan membandingkan secara komprehensif kinerja algoritma K-Nearest Neighbors dan Naive Bayes. Tujuan spesifik dari penelitian ini adalah (1) menghasilkan model klasifikasi pelanggan berdasarkan data transaksi historis, (2) membandingkan performa kedua algoritma untuk menentukan model yang paling efektif, dan (3) menyediakan sebuah sistem pendukung keputusan yang dapat diadopsi oleh perusahaan untuk meningkatkan retensi dan merumuskan strategi pemasaran berbasis data yang lebih tepat sasaran.

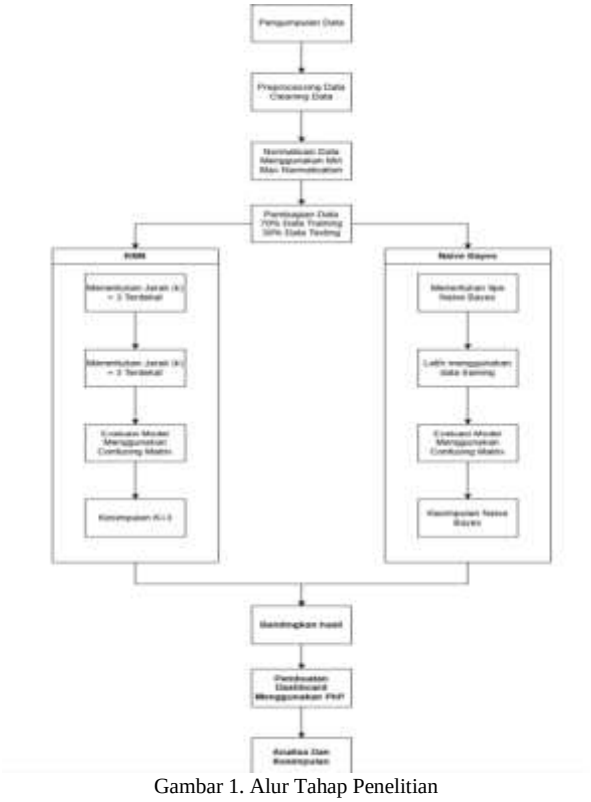
## 2. Metode Penelitian

### 2.1. Landasan Teori

Penelitian ini menerapkan kerangka kerja Knowledge Discovery in Databases (KDD), yaitu sebuah proses untuk menemukan pengetahuan yang valid dan bermanfaat dari kumpulan data dalam jumlah besar. Proses KDD meliputi beberapa tahapan seperti pembersihan data, integrasi, seleksi, transformasi, penambangan data, evaluasi pola, dan presentasi pengetahuan.Inti dari proses KDD adalah Data Mining, di mana teknik statistik dan kecerdasan buatan diterapkan untuk mengeksplorasi dan mengekstraksi pola dari data. Salah satu teknik utama dalam data mining adalah klasifikasi, yaitu proses untuk mengidentifikasi kelas dari sebuah data berdasarkan atribut tertentu dengan mempelajari data latih yang telah memiliki label.

### 2.2. Kerangka Penelitian

Proses penelitian mengikuti alur sistematis yang digambarkan pada Gambar 1. Tahapan dimulai dari pengumpulan data, dilanjutkan dengan pra-pemrosesan, pembagian data, penerapan kedua model (KNN dan Naive Bayes), hingga evaluasi dan perbandingan hasil untuk menarik kesimpulan.




Gambar 1. Alur Tahap Penelitian

### 2.3. Akuisisi dan Karakteristik Dataset

Sumber data yang digunakan adalah data primer dari catatan transaksi penjualan di CV Cahaya Alam Indah. Dari total 10.241 catatan transaksi mentah, dilakukan proses pembersihan data untuk mengatasi data yang tidak relevan dan duplikat, sehingga menghasilkan dataset final sebanyak 220 record yang siap digunakan untuk analisis. Dataset ini terdiri dari tiga atribut utama yang digunakan sebagai fitur, yang dirangkum pada Tabel 1.

DOI: https://doi.org/10.31004/riggs.v4i3.2697

Tabel 1. Atribut yang Digunakan dalam Penelitian

| Atribut | Deskripsi |
|---|---|
| Frekuensi Transaksi | Jumlah total transaksi pembelian oleh pelanggan. |
| Jumlah Pembelian | Akumulasi total galon yang dibeli pelanggan. |
| Preferensi Produk | Jenis produk yang paling sering dibeli. |

## 2.4. Tahap Pra-pemrosesan Data

### 2.4.1. Normalisasi Min-Max

Data numerik dinormalisasi menggunakan metode Min-Max Normalization untuk mengubah rentang nilainya ke dalam skala seragam (0 hingga 1), memastikan setiap fitur memiliki bobot yang seimbang. Rumus yang digunakan adalah sebagai berikut:

$$X\_norm = (X - X\_min) / (X\_max - X\_min) \quad (1)$$

### 2.4.2. Pembagian Dataset

Dataset yang terdiri dari 220 record dibagi dengan rasio 70:30, yaitu 154 record sebagai data latih dan 66 record sebagai data uji.

## 2.5. Implementasi Algoritma

### 2.5.1. K-Nearest Neighbor (KNN)

KNN adalah algoritma berbasis instance yang mengklasifikasikan data baru berdasarkan mayoritas kelas dari K tetangga terdekatnya. Jarak antar data dihitung menggunakan metrik Euclidean Distance (rumus 2). Penelitian ini menetapkan nilai k=3.

$$d(x,y) = SQRT(SUM((x\_i - y\_i)^2)) \quad (2)$$

### 2.5.2. Naive Bayes

Naive Bayes adalah metode klasifikasi probabilistik berdasarkan Teorema Bayes, yang bekerja dengan asumsi bahwa setiap fitur bersifat independen. Probabilitas posterior P(H|E) dari suatu kelas dihitung menggunakan rumus (3):

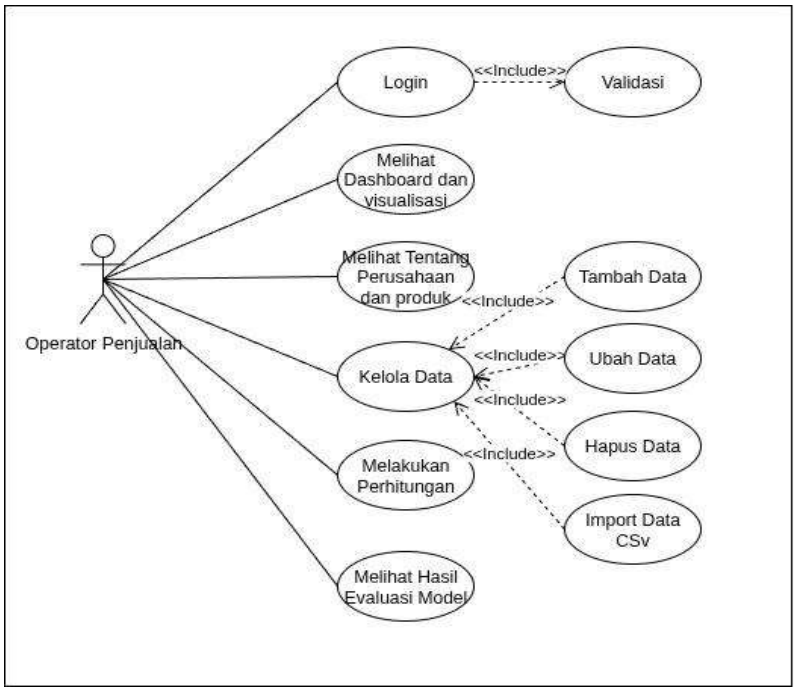
$$P(H|E) = (P(E|H)\ \ P(H)) / P(E) \quad (3)$$

## 2.6. Perancangan Sistem

Perancangan sistem yang matang adalah kunci implementasi model yang berhasil. Untuk mencapai ini, kita menggunakan Unified Modeling Language (UML), sebuah bahasa visual standar yang menyederhanakan kompleksitas sistem perangkat lunak. UML bukan sekadar alat, melainkan jembatan komunikasi antara pengembang, analis, dan pemangku kepentingan, memastikan semua pihak memiliki pemahaman yang sama tentang proyek.

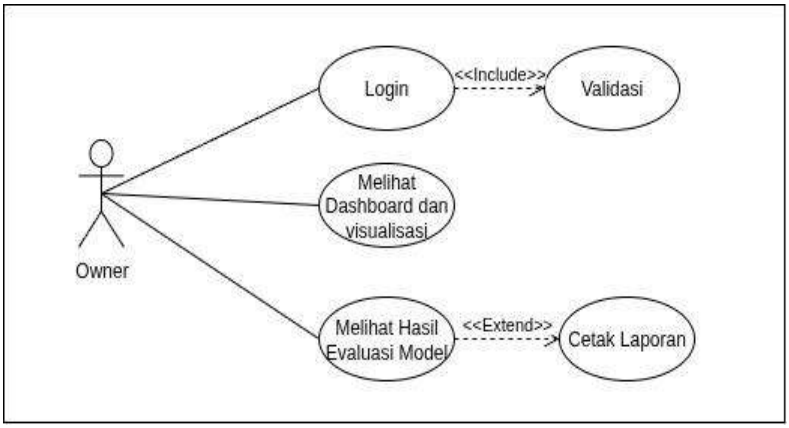
UML menyediakan berbagai diagram untuk memodelkan aspek berbeda dari sistem. Diagram Use Case menjadi fokus utama untuk memodelkan interaksi antara pengguna dan sistem. Dalam kasus ini, aktor utamanya adalah Owner dan Operator. Diagram ini memetakan fungsionalitas sistem dari sudut pandang mereka. Owner mungkin dapat memantau kinerja atau mengelola hak akses, sedangkan Operator dapat memasukkan data atau menjalankan laporan.

DOI: https://doi.org/10.31004/riggs.v4i3.2697




Gambar 2. Diagram Usecase Operator Penjualan

Pada Gambar 2 menunjukkan peran Operator Penjualan dalam sistem, yang memiliki akses ke berbagai fitur seperti login, pengelolaan data (konsumen dan transaksi), serta fungsi klasifikasi dan evaluasi model machine learning (KNN dan Naive Bayes). Operator juga dapat melihat hasil klasifikasi, evaluasi model melalui confusion matrix, dan visualisasi berbentuk diagram batang.




Gambar 3. Diagram Usecase Owner

Pada Gambar 3 menggambarkan peran Owner yang mengakses sistem melalui dashboard. Owner memiliki akses untuk melihat hasil klasifikasi (menggunakan KNN dan Naive Bayes), evaluasi model melalui confusion matrix, metrik evaluasi (akurasi, presisi, recall, dan F1 score), serta visualisasi dalam bentuk diagram batang. Selain itu, Owner juga dapat mencetak seluruh hasil klasifikasi dan evaluasi. Diagram ini menekankan pada peran Owner sebagai pihak pemantau hasil sistem secara menyeluruh tanpa terlibat dalam input data.

2.7. Metrik Evaluasi Kinerja

Evaluasi performa kedua model dilakukan dengan membandingkan hasil prediksi pada data uji dengan kelas aktualnya menggunakan Confusion Matrix. Dari matriks tersebut, metrik kuantitatif utama yang dihitung adalah Akurasi, Presisi, Recall, dan F1-Score. Akurasi, yang mengukur rasio prediksi benar secara keseluruhan, dihitung dengan rumus (4):

$$\text{Akurasi} = (TP+TN) / (TP+TN+FP+FN) \quad (4)$$

DOI: https://doi.org/10.31004/riggs.v4i3.2697

## 3. Hasil dan Diskusi

### 3.1. Hasil Perhitungan Manual

Untuk memberikan pemahaman transparan, berikut adalah contoh perhitungan untuk satu data uji, "SMK Al-Jabar" (fitur ternormalisasi: Frekuensi=0.231, Jumlah=0.280, Produk=1.0).

#### 3.1.1. Perhitungan K-Nearest Neighbor (KNN)

Jarak Euclidean dihitung antara data uji dengan seluruh 154 data latih (Tabel 2). Tiga tetangga terdekat (k=3) kemudian diidentifikasi (Tabel 3) untuk menentukan kelas mayoritas.

Tabel 2. Contoh Perhitungan Jarak Euclidean

| No | Nama Konsumen | Jarak |
|---|---|---|
| 1 | Aan | 0.4675 |
| .. | ... | ... |
| 154 | SMK Al-Ikhlas | 0.0885 |

Tabel 3. Tiga Tetangga Terdekat (k=3)

| No | Tetangga Terdekat | Klasifikasi |
|---|---|---|
| 1 | Nci Babakan | Cenderung Loyal |
| 2 | Juhana | Cenderung Loyal |
| 3 | Salim | Cenderung Loyal |

Hasil prediksi KNN adalah Cenderung Loyal.

#### 3.1.2. Perhitungan Naive Bayes

Probabilitas dihitung berdasarkan nilai mean, deviasi standar pada Tabel 4, dan probabilitas prior dari data latih. Kelas dengan probabilitas posterior tertinggi menjadi prediksi. Untuk data uji ini, hasilnya adalah Cenderung Loyal.

Tabel 4. Nilai Mean dan Deviasi Standar

| Kelas | Fitur | Mean (μ) | Deviasi Standar (σ) |
|---|---|---|---|
| Loyal | Frekuensi | 0,7 | 0,2 |
| Loyal | Jumlah Pembelian | 0,9 | 0,1 |
| Loyal | Produk | 0,7 | 0,3 |
| Cenderung Loyal | Frekuensi | 0,4 | 0,2 |
| Cenderung Loyal | Jumlah Pembelian | 0,5 | 0,2 |
| Cenderung Loyal | Produk | 0,7 | 0,2 |
| Tidak Loyal | Frekuensi | 0,2 | 0,1 |
| Tidak Loyal | Jumlah Pembelian | 0,3 | 0,1 |
| Tidak Loyal | Produk | 0,4 | 0,2 |

Kelas dengan probabilitas posterior tertinggi menjadi prediksi. Untuk data uji ini, hasilnya adalah Cenderung Loyal. berikut adalah hasil perhitungan propabilitasnya posteriornya pada tabel 5.

Tabel 5. Contoh Perhitungan Probabilitas Posterior

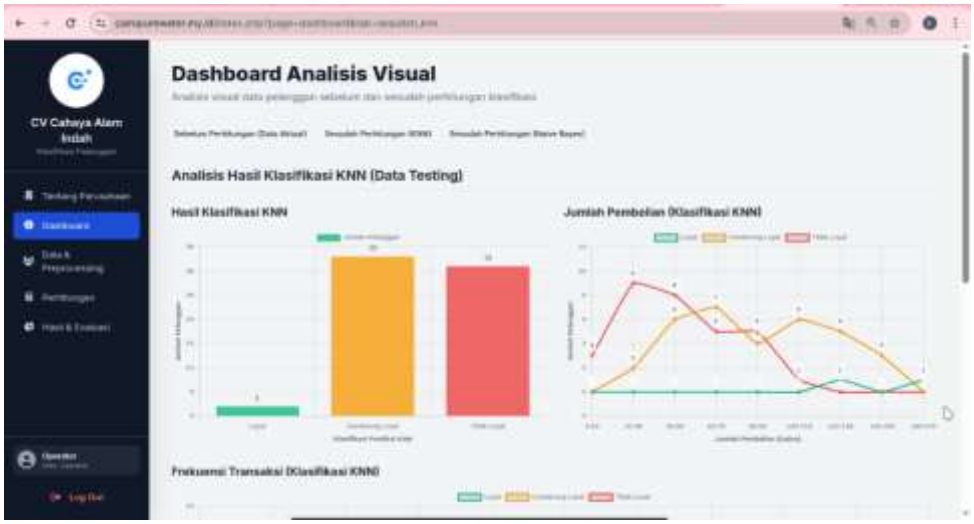
| Probabilitas | Fitur | Total nilai Log |
|---|---|---|
| Loyal | Frekuensi | |
| Loyal | Jumlah Pembelian | -35,4740989 |
| Loyal | Produk | |
| Cenderung Loyal | Frekuensi | |
| Cenderung Loyal | Jumlah Pembelian | -0,2672444651 |

DOI: https://doi.org/10.31004/riggs.v4i3.2697

| | | |
|---|---|---|
| Cenderung Loyal | Produk | |
| Tidak Loyal | Frekuensi | |
| Tidak Loyal | Jumlah Pembelian | -1,361518154 |
| Tidak Loyal | Produk | |

3.2. Hasil Implementasi Sistem

Logika perhitungan di atas diimplementasikan ke dalam sistem informasi SIKAPE. Hasil pengujian Black Box menunjukkan seluruh fungsionalitas sistem berjalan sesuai harapan. Sistem menyajikan hasil analisis dalam dashboard visual pada Gambar 5 dan halaman evaluasi pada Gambar 6.




Gambar 5. Dashboard Visual

Dashboard ini menyajikan analisis visual yang komprehensif dari hasil klasifikasi pelanggan menggunakan metode K-Nearest Neighbors (KNN) pada data uji. Secara ringkas, dashboard ini menampilkan distribusi pelanggan ke dalam tiga kategori loyalitas Loyal, Cenderung Loyal, dan Tidak Loyal melalui grafik batang


Gambar 6. Halaman Evaluasi

Dashboard ini menampilkan perbandingan hasil dan evaluasi kinerja dua model klasifikasi, yaitu K-Nearest Neighbors (KNN) dan Naive Bayes, berdasarkan 66 data uji. Halaman ini menyajikan metrik evaluasi kunci seperti akurasi, presisi, recall, dan F1-Score untuk setiap kelas loyalitas pelanggan (Loyal, Cenderung Loyal, dan Tidak Loyal). Di bagian bawah, terdapat confusion matrix untuk setiap model, yang secara visual merangkum performa klasifikasi dengan menunjukkan jumlah prediksi yang benar dan salah. Tujuannya adalah untuk memudahkan perbandingan kinerja kedua algoritma dan menentukan mana yang lebih efektif.

DOI: https://doi.org/10.31004/riggs.v4i3.2697

### 3.3. Evaluasi Kinerja Model

Evaluasi kuantitatif dilakukan pada 66 data uji untuk mengukur kinerja model. Hasilnya disajikan dalam confusion matrix pada Tabel 6 dan Tabel 7, yang menggambarkan ringkasan performa klasifikasi. Selanjutnya, perbandingan metrik evaluasi seperti akurasi, presisi, dan recall dirangkum pada Tabel 8 untuk memberikan penilaian yang lebih komprehensif.

Tabel 6. Confusion Matrix KNN

| KNN | | Prediksi | | |
|---|---|---|---|---|
| Aktual | | Loyal | C. Loyal | T. Loyal |
| | Loyal | 1 | 5 | 0 |
| | C. Loyal | 1 | 26 | 3 |
| | T. Loyal | 0 | 2 | 28 |

Tabel 7. Confusion Matrix Naive Bayes

| Naive Bayes | | Prediksi | | |
|---|---|---|---|---|
| Aktual | | Loyal | C. Loyal | T. Loyal |
| | Loyal | 1 | 5 | 0 |
| | C. Loyal | 2 | 24 | 3 |
| | T. Loyal | 0 | 3 | 28 |

Tabel 8. Perbandingan Metrik Evaluasi Kinerja Model

| Metrik | Kelas | Naive Bayes | KNN |
|---|---|---|---|
| Akurasi | Keseluruhan | 80.3% | 83.3% |
| Presisi | Loyal | 33.3% | 50.0% |
| | C. Loyal | 75.0% | 78.8% |
| | T. Loyal | 90.3% | 90.3% |
| Recall | Loyal | 16.7% | 16.7% |
| | C. Loyal | 82.8% | 86.7% |
| | T. Loyal | 90.3% | 93.3% |
| F1-Score | Loyal | 22.2% | 25.0% |
| | C. Loyal | 78.7% | 82.5% |
| | T. Loyal | 90.3% | 91.8% |

### 3.4. Pembahasan

Analisis terhadap hasil pada Tabel 7 secara definitif menunjukkan bahwa algoritma KNN memiliki kinerja yang lebih unggul dibandingkan Naive Bayes untuk studi kasus ini. Keunggulan ini tidak hanya tecermin pada akurasi keseluruhan yang lebih tinggi (83.3% vs 80.3%), tetapi juga pada nilai F1-Score yang secara konsisten lebih superior di semua kelas.Salah satu temuan paling signifikan dari penelitian ini adalah kemampuan tinggi kedua model, terutama KNN, dalam mengidentifikasi pelanggan "Tidak Loyal". Dengan F1-Score mencapai 91.8%, model ini sangat andal dalam mendeteksi pelanggan yang berisiko berhenti berlangganan.Meskipun demikian, penelitian ini juga mengungkap sebuah keterbatasan fundamental pada kedua model, yaitu kinerja yang sangat buruk dalam mengidentifikasi kelas "Loyal". Nilai recall sebesar 16.7% menunjukkan bahwa model hanya berhasil menemukan 1 dari 6 pelanggan yang sebenarnya loyal. Akar penyebab dari masalah ini adalah ketidakseimbangan data (class imbalance) yang parah dalam dataset pelatihan. Keterbatasan ini menyoroti dampak langsung dari proses pencatatan data manual yang tidak terstruktur.

DOI: https://doi.org/10.31004/riggs.v4i3.2697

## 4. Kesimpulan

Berdasarkan hasil analisis dan pembahasan, dapat ditarik beberapa kesimpulan utama. Pertama, penelitian ini telah berhasil merancang, mengimplementasikan, dan menguji sebuah sistem fungsional untuk klasifikasi loyalitas pelanggan. Kedua, hasil evaluasi kuantitatif secara jelas menunjukkan bahwa algoritma KNN memiliki kinerja yang lebih unggul dibandingkan dengan Naive Bayes, dengan mencapai tingkat akurasi keseluruhan sebesar 83.3%. Ketiga, terlepas dari keunggulannya, penelitian ini mengidentifikasi adanya keterbatasan signifikan pada model, yaitu kesulitan dalam mengidentifikasi pelanggan dari kelas "Loyal", yang disebabkan oleh masalah ketidakseimbangan data yang parah pada dataset pelatihan. Beberapa saran yang dapat diberikan antara lain, untuk implementasi bisnis: 1. Adopsi Model KNN: CV Cahaya Alam Indah disarankan untuk mengadopsi model KNN sebagai alat bantu pengambilan keputusan strategis untuk segmentasi pelanggan, 2. Fokus pada Retensi: Perusahaan dapat memanfaatkan kekuatan model untuk secara akurat mengidentifikasi pelanggan "Tidak Loyal" dan menargetkan mereka dengan program retensi khusus, 3. Perbaikan Proses Pengumpulan Data: Sangat penting bagi perusahaan untuk mulai beralih ke sistem pencatatan transaksi digital yang terstruktur untuk membangun dataset yang lebih seimbang di masa depan. Untuk penelitian selanjutnya: 1. Penanganan Class Imbalance: Disarankan untuk menerapkan teknik penanganan data tidak seimbang, seperti oversampling menggunakan metode SMOTE, untuk meningkatkan kemampuan model dalam mengenali kelas minoritas, 2. Eksplorasi Algoritma Lanjutan: Penelitian mendatang dapat mengeksplorasi algoritma klasifikasi yang lebih kompleks seperti Random Forest atau SVM, yang dikenal memiliki ketahanan lebih baik terhadap data yang tidak seimbang, 3. Feature Engineering: Disarankan untuk melakukan rekayasa fitur dengan menambahkan atribut baru di luar data transaksi, seperti lama waktu menjadi pelanggan (customer tenure), yang berpotensi meningkatkan daya prediksi model.

## Referensi

[1] Gultom, D. K., Arif, M., & Fahmi, M. (2020). Determinasi Kepuasan Pelanggan Terhadap Loyalitas Pelanggan Melalui Kepercayaan. Maneggio : Jurnal Ilmiah Magister Manajemen, 3(2), pp. 171–180.
[2] I. N. Suwandana dan I. K. A. Purnamawati, "Analisis Kualitas Pelayanan dan Kepuasan Pelanggan pada Depot Air Minum Isi Ulang," *Jurnal Ilmiah Manajemen dan Bisnis*, vol. 9, no. 2, pp. 150–159, 2021.
[3] Kementerian Lingkungan Hidup dan Kehutanan Republik Indonesia, *Status Lingkungan Hidup Indonesia Tahun 2020*, Jakarta: KLHK, 2021.
[4] Badan Pusat Statistik, *Statistik Lingkungan Hidup Indonesia 2021*, Jakarta: BPS, 2021.
[5] Badan Nasional Penanggulangan Bencana (BNPB), *Laporan Tahunan Bencana Indonesia 2022*, Jakarta: BNPB, 2023.
[6] Kementerian PUPR, *Rencana Strategis Direktorat Jenderal Cipta Karya 2020–2024*, Jakarta: KemenPUPR, 2020.
[7] Kementerian Pekerjaan Umum dan Perumahan Rakyat, *Kajian Strategi Pengembangan Sistem Penyediaan Air Minum di Wilayah Perkotaan Tahun 2021*, Jakarta: Direktorat Jenderal Cipta Karya, 2021.
[8] Peningkatan Produk dengan Data Pelanggan, Digima, 2024. [Online]. Tersedia: https://digima.co.id/peningkatan-produk-dengan-data-pelanggan/. [Diakses: 12-Mei-2025].
[9] Segmentasi Transaksional (Analisis RFM): Pengertian, Strategi, dan Contohnya, Kledo, 2023. [Online]. Tersedia: https://kledo.com/blog/segmentasi-transaksional/. [Diakses: 12-Mei-2025].
[10] Cara Meningkatkan Penjualan Berdasarkan Data Pelanggan," PTNAS, 2024. [Online]. Tersedia: https://ptnas.co.id/blog/cara-meningkatkan-penjualan-bisnis/. [Diakses: 12-Mei-2025].
[11] R. Rahmadana et al., "Penerapan Algoritma KNN dalam Analisis Kepuasan Pelanggan pada PDAM Batiwakkal Berau," *Jurnal Teknologi dan Sistem Informasi*, vol. 12, no. 1, pp. 50-62, 2024.
[12] B. Rekayasa et al., "Perbandingan Algoritma SVM dan Naïve Bayes dalam Klasifikasi Kualitas Air," *Jurnal Sistem Informasi*, vol. 15, no. 3, pp. 120-130, 2025.
[13] E. F. Wati, H. Ginting, dan F. W. Pratiwi, "Klasifikasi loyalitas pelanggan menggunakan metode Naïve Bayes, C4.5 dan KNN," Indonesian Journal of Information System and Technology (IJISTECH), vol. 4, no. 1, pp. 30–36, 2024.
[14] R. Y. Hayuningtyas, "Klasifikasi Loyalitas Pelanggan Menggunakan Naïve Bayes pada Konsumen Starbucks," Jurnal Teknik dan Komputer, vol. 5, no. 1, pp. 45–51, 2024.
[15] M. A. Alif, "Penerapan Algoritma Naive Bayes untuk Klasifikasi Data Mahasiswa," Jurnal Ilmiah Teknologi Informasi, vol. 6, no. 2, pp. 55–60, 2020.

DOI: https://doi.org/10.31004/riggs.v4i3.2697