Terakreditasi Peringkat 4
Keputusan Direktur Jenderal Pendidikan Tinggi, Riset, dan Teknologi, No. 225/E/KPT/2022
Masa berlaku dari Volume 8 Nomor 1 Tahun 2022 sampai dengan Volume 12 Nomor 2 Tahun 2026

Published online at: **https://ejournal.raharja.ac.id/**

JOURNAL CERITA:
Creative Education of Research in Information Technology and Artificial informatics

Vol. 12 No. 2 (2026) 162-177 e-ISSN: 2655 - 2574

Perbandingan Algoritma Naïve Bayes dan KNN pada Analisis Sentimen Kepuasan Pelanggan E-Commerce

Wahyu Hidayat[*1], Fahri Shaftian[2], Emelia Lopez[3], Ajay Supriyadi[4], Riyan Hermawan[4]
[14]Sistem Informasi, Fakultas Sains dan Teknologi, Universitas Raharja, Tangerang, Indonesia
[2]Magister Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Raharja, Tangerang, Indonesia
[3]Teknik Informatika, Instituto Profissional de Canossa (IPDC) Dili, Timor-Leste
[5]Teknik Informatika, Fakultas Ilmu Komputer, Universitas Pamulang, Tangerang, Indonesia
Email: **wahyu@raharja.info*[1]**; fahri.shaftian@raharja.info[2]; emelia@gmail.com[3]; ajay.supriadi@raharja.info[4]; riyanhermawan727@gmail.com[5]

Hidayat, W., Shaftian, F., Lopez, E., Supriyadi, A., & Hermawan, R. (2026). Perbandingan Algoritma Naïve Bayes dan KNN pada Analisis Sentimen Kepuasan Pelanggan E-Commerce. *Journal Cerita: Creative Education of Research in Information Technology and Artificial Informatics*, 12(2), 162-177

DOI: https://doi.org/10.33050/cerita.v12i2.4434

ABSTRAK

Pertumbuhan e-commerce di Indonesia yang semakin pesat menyebabkan peningkatan jumlah ulasan pelanggan pada platform digital. Ulasan pelanggan merupakan sumber informasi penting bagi perusahaan dalam memahami tingkat kepuasan pelanggan terhadap produk maupun layanan yang diberikan. Namun, volume data ulasan yang besar serta karakteristik teks yang tidak terstruktur menyebabkan proses evaluasi manual menjadi kurang efisien dan rentan terhadap subjektivitas. Oleh karena itu, diperlukan pendekatan berbasis machine learning untuk melakukan klasifikasi sentimen secara otomatis. Penelitian ini bertujuan membandingkan performa algoritma Naïve Bayes dan K-Nearest Neighbor (KNN) dalam analisis sentimen kepuasan pelanggan e-commerce pada PT XYZ. Kebaruan penelitian ini terletak pada evaluasi komparatif kedua algoritma menggunakan dataset ulasan pelanggan internal PT XYZ yang direpresentasikan dengan pembobotan TF-IDF serta dianalisis pada klasifikasi sentimen multikelas (positif, negatif, dan netral). Dataset penelitian terdiri atas 3.500 ulasan pelanggan yang dibagi menggunakan metode Split Validation dengan rasio 80:20, yaitu 2.800 data pelatihan dan 700 data pengujian. Tahapan penelitian meliputi text preprocessing berupa cleaning, case folding, tokenizing, stopword removal, dan stemming, dilanjutkan dengan pembobotan fitur menggunakan Term Frequency–Inverse Document Frequency (TF-IDF). Proses klasifikasi dilakukan menggunakan algoritma Naïve Bayes dan KNN (K = 5), sedangkan evaluasi model menggunakan metrik accuracy, precision, recall, dan F1-score. Hasil penelitian menunjukkan bahwa algoritma K-Nearest Neighbor (KNN) menghasilkan performa terbaik dengan nilai accuracy 92,57%, precision 91,56%, recall 92,73%, dan F1-score 92,14%, sedangkan algoritma Naïve Bayes memperoleh accuracy 86,85%, precision 86,30%, recall 88,50%, dan F1-score 87,30%. Hasil tersebut menunjukkan bahwa pendekatan berbasis kemiripan (similarity-based classification) pada KNN lebih efektif dibandingkan pendekatan probabilistik Naïve Bayes dalam mengklasifikasikan sentimen pelanggan pada dataset

Received: 20-02-2026 | Accepted: 20-06-2026 | Published Online: 08-08-2026

penelitian ini. Temuan penelitian ini diharapkan dapat menjadi referensi dalam pengembangan sistem analisis sentimen untuk mendukung evaluasi kualitas layanan dan pengambilan keputusan berbasis data pada perusahaan e-commerce.

Kata kunci: Analisis Sentimen, Naïve Bayes, K-Nearest Neighbor (KNN), E-Commerce, TF-IDF, Kepuasan Pelanggan

ABSTRACT

The rapid growth of e-commerce in Indonesia has led to a significant increase in the volume of customer reviews on digital platforms. Customer reviews provide valuable information for companies to understand customer satisfaction with products and services. However, the large volume of reviews and the unstructured nature of textual data make manual evaluation inefficient and prone to subjectivity. Therefore, machine learning approaches are required to perform automatic sentiment classification. This study aims to compare the performance of the Naïve Bayes and K-Nearest Neighbor (KNN) algorithms for customer satisfaction sentiment analysis in the e-commerce platform of PT XYZ. The dataset consists of 3,500 customer reviews, divided into 80% training data (2,800 reviews) and 20% testing data (700 reviews) using the Split Validation method. The preprocessing stage includes cleaning, case folding, tokenization, stopword removal, and stemming, followed by feature weighting using the Term Frequency–Inverse Document Frequency (TF-IDF) method. The classification models were implemented using RapidMiner Studio 10.x, while model performance was evaluated using accuracy, precision, recall, and F1-score metrics. The experimental results indicate that the K-Nearest Neighbor (KNN) algorithm outperformed Naïve Bayes, achieving an accuracy of 92.57%, precision of 91.56%, recall of 92.73%, and F1-score of 92.14%, whereas Naïve Bayes achieved an accuracy of 86.85%, precision of 86.30%, recall of 88.50%, and F1-score of 87.30%. These findings demonstrate that KNN is more effective in classifying customer sentiment on the PT XYZ e-commerce review dataset. The novelty of this study lies in the comparative evaluation of Naïve Bayes and KNN using an internal real-world customer review dataset from PT XYZ with TF-IDF feature weighting and multiclass sentiment classification, providing empirical evidence for selecting an appropriate sentiment classification model in e-commerce environments.

Keywords: *Sentiment Analysis, Naïve Bayes, K-Nearest Neighbor (KNN), E-Commerce, TF-IDF, Customer Satisfaction*

I. PENDAHULUAN

Perkembangan teknologi informasi dan komunikasi telah mendorong transformasi signifikan dalam berbagai sektor industri, termasuk sektor perdagangan elektronik (e-commerce). Dalam beberapa tahun terakhir, pertumbuhan e-commerce di Indonesia mengalami peningkatan yang sangat pesat seiring meningkatnya penetrasi internet, penggunaan perangkat digital, serta perubahan perilaku masyarakat menuju transaksi berbasis daring. Kehadiran e-commerce tidak hanya mengubah pola transaksi konvensional menjadi digital, tetapi juga memberikan dampak terhadap pertumbuhan ekonomi nasional melalui peningkatan aktivitas perdagangan, efisiensi distribusi, serta perluasan akses pasar. Penelitian yang dilakukan oleh Nopiah et al. menunjukkan bahwa perkembangan e-commerce memberikan kontribusi positif terhadap pertumbuhan ekonomi Indonesia, khususnya melalui integrasi teknologi keuangan (financial technology) yang mempercepat aktivitas ekonomi digital dan meningkatkan produktivitas sektor perdagangan (Nopiah et al., 2024). Selain itu, penelitian Situmorang dan Prima menjelaskan bahwa pertumbuhan transaksi e-commerce di Indonesia memiliki pengaruh terhadap peningkatan Produk Domestik Regional Bruto (PDRB), sehingga memperlihatkan pentingnya sektor perdagangan digital sebagai salah satu pendorong pembangunan ekonomi nasional (Situmorang & Prima, 2025).

Seiring meningkatnya jumlah transaksi digital, volume data ulasan pelanggan pada platform e-commerce juga mengalami peningkatan yang sangat besar. Data ulasan pelanggan menjadi sumber informasi penting bagi perusahaan karena memuat berbagai opini,

pengalaman, serta tingkat kepuasan pelanggan terhadap produk maupun layanan yang diterima. Ulasan pelanggan umumnya berbentuk data teks tidak terstruktur (unstructured text data) dengan karakteristik bahasa yang kompleks, seperti penggunaan bahasa informal, singkatan, kesalahan penulisan, campuran bahasa, serta ekspresi emosional yang beragam. Kondisi tersebut menyebabkan proses evaluasi manual terhadap ulasan pelanggan menjadi tidak efisien, memerlukan waktu yang lama, serta rentan terhadap subjektivitas manusia. Oleh karena itu, perusahaan memerlukan pendekatan berbasis teknologi yang mampu mengolah data ulasan pelanggan secara otomatis guna memperoleh informasi yang lebih akurat terkait tingkat kepuasan pelanggan dan kualitas layanan yang diberikan (Zahra et al., 2026).

Salah satu pendekatan yang banyak digunakan untuk memahami persepsi pelanggan adalah analisis sentimen (sentiment analysis). Analisis sentimen merupakan cabang dari Natural Language Processing (NLP) yang bertujuan untuk mengidentifikasi, mengekstraksi, dan mengklasifikasikan opini atau emosi pengguna terhadap suatu objek ke dalam kategori tertentu, seperti sentimen positif, negatif, atau netral. Implementasi analisis sentimen pada sektor e-commerce menjadi semakin penting karena mampu membantu perusahaan memahami kebutuhan pelanggan secara lebih cepat dan efisien melalui pemrosesan data teks dalam jumlah besar. Selain itu, hasil analisis sentimen dapat digunakan untuk mendukung pengambilan keputusan bisnis, meningkatkan kualitas pelayanan, mengoptimalkan strategi pemasaran, serta memperbaiki pengalaman pelanggan (customer experience) berdasarkan persepsi pengguna terhadap layanan perusahaan (Hasibuan et al., 2024; Dikananda et al., 2024).

Dalam implementasinya, berbagai metode machine learning telah digunakan untuk melakukan klasifikasi sentimen, di antaranya Naïve Bayes, K-Nearest Neighbor (KNN), Support Vector Machine (SVM), dan Random Forest. Masing-masing algoritma memiliki karakteristik, keunggulan, serta tingkat akurasi yang berbeda dalam mengolah data teks. Algoritma Naïve Bayes dan K-Nearest Neighbor dipilih dalam penelitian ini karena keduanya merupakan metode klasifikasi yang banyak digunakan pada analisis sentimen berbasis teks serta memiliki karakteristik yang berbeda. Naïve Bayes menggunakan pendekatan probabilistik dengan asumsi independensi antarfitur, sedangkan KNN menerapkan pendekatan berbasis kemiripan (similarity-based classification) yang mengklasifikasikan data berdasarkan kedekatan antar dokumen. Perbedaan karakteristik tersebut menjadikan kedua algoritma menarik untuk dibandingkan dalam mengidentifikasi metode yang paling sesuai untuk klasifikasi sentimen pada dataset ulasan pelanggan PT XYZ. Penelitian oleh Abdillah et al. (2024) menunjukkan bahwa algoritma Naïve Bayes dan KNN sama-sama mampu digunakan untuk klasifikasi sentimen digital, dengan perbedaan performa yang dipengaruhi oleh karakteristik dataset, distribusi kelas, serta tahapan preprocessing yang diterapkan. Hasil penelitian tersebut menunjukkan bahwa kedua algoritma masih relevan dan layak digunakan dalam penelitian analisis sentimen berbasis ulasan pengguna (Abdillah et al., 2024). Penelitian lain juga menunjukkan bahwa model berbasis Naïve Bayes masih menjadi salah satu algoritma yang efektif untuk klasifikasi sentimen pada ulasan produk digital karena memiliki kemampuan menangani data berdimensi tinggi secara efisien(Zahra et al., 2026).

Algoritma Naïve Bayes merupakan metode klasifikasi probabilistik yang bekerja berdasarkan Teorema Bayes dengan asumsi independensi antarfitur. Metode ini dikenal memiliki keunggulan dalam proses klasifikasi data teks karena mampu bekerja secara cepat, sederhana, serta memiliki performa yang stabil pada data berdimensi tinggi. Penelitian oleh Berlianti dan Hidayat (2024) menunjukkan bahwa penerapan Naïve Bayes Classifier pada analisis sentimen ulasan produk kecantikan mampu menghasilkan akurasi sebesar 82,77% setelah optimasi hyperparameter, sehingga efektif digunakan untuk memahami preferensi dan opini konsumen(Berlianti & Hidayat, 2024). Selain itu, penelitian Riki et al. (2024) pada ulasan produk e-commerce Shopee menunjukkan bahwa algoritma Naïve Bayes mampu mengklasifikasikan sentimen pengguna secara efektif pada data ulasan dalam jumlah besar, sehingga dapat membantu perusahaan memperoleh informasi mengenai tingkat kepuasan pelanggan dan kualitas layanan secara lebih efisien(Riki et al., 2025).

Di sisi lain, algoritma K-Nearest Neighbor (KNN) bekerja menggunakan pendekatan

berbasis kedekatan (similarity-based classification) yang mengklasifikasikan data baru berdasarkan mayoritas kelas dari tetangga terdekatnya. Algoritma ini memiliki kemampuan menghasilkan klasifikasi yang kompetitif karena mempertimbangkan kemiripan antar data menggunakan ukuran jarak tertentu. Namun demikian, performa KNN dipengaruhi oleh pemilihan parameter nilai k, metode ekstraksi fitur, serta karakteristik dataset yang digunakan. Penelitian Rayhan et al. (2024) menunjukkan bahwa penerapan algoritma KNN pada analisis sentimen pengguna aplikasi DANA dengan pembobotan TF-IDF mampu menghasilkan performa klasifikasi yang baik dalam mengidentifikasi opini pengguna aplikasi keuangan digital (Rayhan et al., 2024). Penelitian Latifurrohman juga menunjukkan bahwa kombinasi metode KNN dan Naïve Bayes pada analisis sentimen ulasan objek wisata berbasis Google Maps menghasilkan performa klasifikasi yang berbeda tergantung karakteristik data dan proses preprocessing yang diterapkan (Latifurrohman et al., 2025).

Meskipun demikian, hasil penelitian sebelumnya menunjukkan adanya perbedaan performa antara algoritma Naïve Bayes dan KNN pada berbagai kasus analisis sentimen. Penelitian Pratmanto et al. menunjukkan bahwa algoritma Naïve Bayes menghasilkan performa klasifikasi yang lebih baik dibandingkan KNN dalam analisis sentimen ulasan pengguna aplikasi Vidio di Google Play Store (Pratmanto et al., 2024). Sebaliknya, penelitian Muslim et al. (2024) menunjukkan bahwa KNN menghasilkan tingkat akurasi yang lebih tinggi dibandingkan Naïve Bayes pada analisis sentimen ulasan pengguna aplikasi CapCut (Muslim et al., 2024). Perbedaan hasil tersebut menunjukkan bahwa efektivitas algoritma klasifikasi sangat dipengaruhi oleh karakteristik dataset, proses preprocessing, teknik ekstraksi fitur, serta metode pembobotan yang digunakan selama proses analisis sentimen.

Meskipun berbagai penelitian telah membandingkan algoritma Naïve Bayes dan K-Nearest Neighbor pada analisis sentimen, sebagian besar penelitian masih menggunakan dataset publik yang berasal dari Google Play Store, marketplace nasional, maupun aplikasi digital tertentu. Penelitian yang memanfaatkan dataset ulasan pelanggan internal perusahaan e-commerce dengan karakteristik pelanggan yang lebih spesifik masih relatif terbatas. Perbedaan karakteristik dataset tersebut berpotensi memengaruhi performa algoritma klasifikasi sehingga hasil penelitian terdahulu belum tentu dapat digeneralisasikan pada lingkungan perusahaan e-commerce yang berbeda. Oleh karena itu, masih terdapat research gap terkait pemilihan algoritma yang paling efektif untuk analisis sentimen menggunakan dataset internal perusahaan. Selain itu, penelitian Hasbi dan Putro (2025) menunjukkan bahwa teknik pembobotan fitur seperti TF-IDF juga berpengaruh terhadap kualitas representasi data teks dan hasil klasifikasi sentimen.(Hasbi & Putro, 2025).

PT XYZ sebagai perusahaan yang bergerak di bidang e-commerce menghadapi tantangan dalam memahami persepsi pelanggan terhadap layanan yang diberikan. Setiap hari, PT XYZ menerima sejumlah besar ulasan pelanggan terkait kualitas produk, kecepatan pengiriman, harga, pelayanan pelanggan, serta pengalaman penggunaan platform secara keseluruhan. Namun, proses evaluasi terhadap ulasan tersebut masih belum dilakukan secara optimal karena besarnya volume data yang tersedia. Akibatnya, perusahaan mengalami kesulitan dalam memperoleh gambaran sentimen pelanggan secara cepat, objektif, dan komprehensif sebagai dasar pengambilan keputusan bisnis.

Oleh karena itu, penelitian ini bertujuan membandingkan performa algoritma Naïve Bayes dan K-Nearest Neighbor (KNN) untuk menentukan algoritma yang paling efektif dalam mengklasifikasikan sentimen pelanggan pada dataset ulasan internal PT XYZ. Penelitian ini juga mengevaluasi karakteristik kesalahan klasifikasi masing-masing algoritma melalui confusion matrix sehingga diperoleh pemahaman yang lebih komprehensif terhadap performa model.

Kebaruan (novelty) penelitian ini terletak pada evaluasi komparatif algoritma Naïve Bayes dan K-Nearest Neighbor (KNN) menggunakan dataset ulasan pelanggan internal PT XYZ yang belum banyak digunakan pada penelitian sebelumnya. Berbeda dengan sebagian besar penelitian terdahulu yang memanfaatkan dataset publik, penelitian ini menguji performa kedua algoritma pada data riil perusahaan dengan klasifikasi sentimen multikelas (positif, negatif, dan netral). Selain membandingkan nilai accuracy, precision, recall, dan F1-score, penelitian ini juga menganalisis karakteristik kesalahan klasifikasi melalui confusion matrix sehingga memberikan pemahaman yang lebih

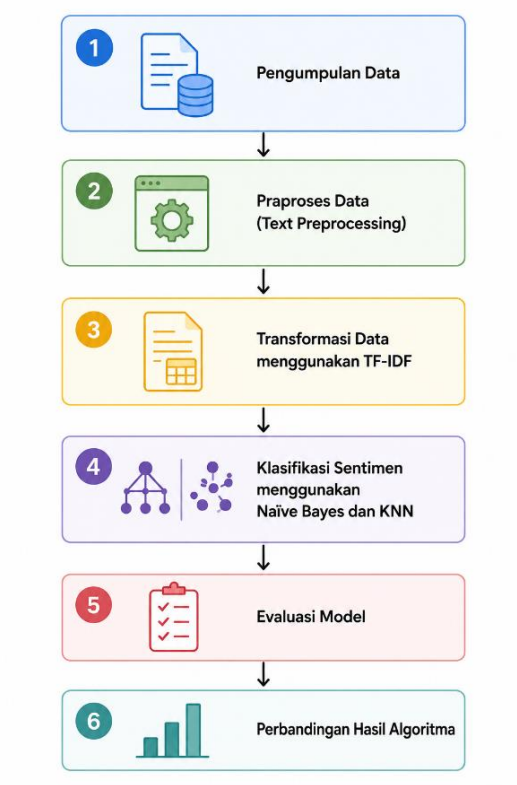
komprehensif mengenai keunggulan dan keterbatasan masing-masing algoritma. Hasil penelitian diharapkan dapat menjadi bukti empiris dalam pemilihan algoritma klasifikasi sentimen yang paling sesuai untuk lingkungan e-commerce serta mendukung pengembangan sistem pengambilan keputusan berbasis data (data-driven decision making).

II. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif eksperimental dengan metode komparatif untuk membandingkan performa algoritma Naïve Bayes dan K-Nearest Neighbor (KNN) dalam analisis sentimen kepuasan pelanggan e-commerce pada PT XYZ. Penelitian diimplementasikan menggunakan RapidMiner Studio versi 10.x sebagai perangkat lunak utama untuk proses praproses data, pembobotan TF-IDF, klasifikasi, dan evaluasi performa model. Eksperimen dijalankan pada komputer dengan sistem operasi Microsoft Windows 11 64-bit, prosesor Intel Core i5/i7, dan memori minimal 8 GB RAM. Operator yang digunakan meliputi Read Excel, Process Documents from Data, TF-IDF, Split Validation, Naïve Bayes, K-Nearest Neighbor (KNN), dan Performance (Classification). Validasi model dilakukan menggunakan metode Split Validation dengan rasio 80:20. Metode ini dipilih karena jumlah dataset relatif besar sehingga mampu menyediakan data pelatihan yang memadai sekaligus mempertahankan data pengujian yang representatif untuk mengevaluasi kemampuan generalisasi model dengan waktu komputasi yang lebih efisien. Tahapan penelitian meliputi pengumpulan data, text preprocessing, pembobotan fitur menggunakan TF-IDF, proses klasifikasi menggunakan algoritma Naïve Bayes dan KNN, serta evaluasi performa model menggunakan metrik accuracy, precision, recall, dan F1-score.(Wiguna & Ginting, 2026; Ni'mah & Arifin, 2020).

A. Tahap Penelitian

Penelitian dilaksanakan melalui beberapa tahapan yang ditunjukkan pada Gambar 1.



Gambar 1. Tahapan Penelitian

Tahapan penelitian pada Gambar 1 dimulai dengan pengumpulan data ulasan pelanggan e-commerce, dilanjutkan dengan praproses data yang meliputi cleaning, case folding, tokenizing, stopword removal, dan stemming. Data yang telah diproses kemudian ditransformasikan menggunakan TF-IDF dan diklasifikasikan menggunakan algoritma Naïve Bayes dan K-Nearest Neighbor (KNN). Selanjutnya, performa model dievaluasi menggunakan metrik accuracy, precision, recall, dan F1-score, kemudian hasil kedua algoritma dibandingkan untuk menentukan model dengan kinerja terbaik.

B. Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan data ulasan pelanggan (customer reviews) pada platform e-commerce PT XYZ. Dataset diperoleh dari sistem internal perusahaan yang berisi opini pelanggan terkait kualitas produk, kecepatan pengiriman, pelayanan pelanggan, harga, serta pengalaman penggunaan platform secara keseluruhan.

Data penelitian berbentuk teks tidak terstruktur (unstructured text) sehingga memerlukan proses pengolahan sebelum dilakukan klasifikasi sentimen. Dataset yang digunakan terdiri atas data ulasan pelanggan yang kemudian diberikan label sentimen ke dalam

kategori positif, negatif, dan netral berdasarkan isi opini pelanggan.

Teknik pengambilan data dilakukan menggunakan metode purposive sampling, yaitu pemilihan data berdasarkan kriteria tertentu, antara lain:

1) Ulasan berasal dari pelanggan aktif PT XYZ.
2) Ulasan memiliki isi teks yang relevan dengan kualitas layanan e-commerce.
3) Ulasan menggunakan bahasa Indonesia.
4) Ulasan tidak mengandung spam atau duplikasi data.

C. Praproses Data (Text Preprocessing)

Tahap text preprocessing dilakukan untuk membersihkan data teks dan meningkatkan kualitas representasi fitur sebelum proses klasifikasi dilakukan. Tahapan preprocessing dalam penelitian ini terdiri atas:

1) Cleaning

Proses cleaning dilakukan untuk menghapus karakter yang tidak diperlukan seperti angka, tanda baca, URL, emoji, simbol, dan karakter khusus lainnya. Menurut penelitian Fahrezi (2024), preprocessing diperlukan untuk menghilangkan berbagai karakter yang tidak relevan agar data dapat digunakan secara optimal pada proses klasifikasi sentimen(Fahrezi & Solichin, 2024).

Tabel 2. Hasil Cleaning

Sebelum Cleaning	Setelah Cleaning
Pengiriman produk sangat cepat 👍 dan kualitas barang memuaskan!!!	Pengiriman produk sangat cepat dan kualitas barang memuaskan
Produk sudah diterima sesuai pesanan ☺	Produk sudah diterima sesuai pesanan
Barang datang terlambat ☹ dan kemasan rusak!!!	Barang datang terlambat dan kemasan rusak

2) Case Folding

Case folding dilakukan dengan mengubah seluruh huruf menjadi huruf kecil (lowercase) agar data memiliki format yang seragam.

Tabel 3. Hasil Case Folding

Sebelum Case Folding	Setelah Case Folding
Pengiriman Produk Sangat Cepat Dan Kualitas Barang Memuaskan	pengiriman produk sangat cepat dan kualitas barang memuaskan
Produk Sudah Diterima Sesuai Pesanan	produk sudah diterima sesuai pesanan
Barang Datang Terlambat Dan Kemasan Rusak	barang datang terlambat dan kemasan rusak

3) Tokenizing

Tahap ini dilakukan untuk memisahkan kalimat menjadi unit kata (token).

Tabel 4. Hasil Tokenisasi

Setelah Normalisasi	Setelah Tokenisasi
pengiriman produk sangat cepat dan kualitas barang memuaskan	[pengiriman, produk, sangat, cepat, dan, kualitas, barang, memuaskan]
produk sudah diterima sesuai pesanan	[produk, sudah, diterima, sesuai, pesanan]
barang datang terlambat dan kemasan rusak	[barang, datang, terlambat, dan, kemasan, rusak]

4) Stopword Removal

Tahap ini dilakukan dengan menghapus kata-kata umum yang tidak memiliki makna signifikan dalam analisis sentimen, seperti dan, yang, di, ke, dan sebagainya.

Tabel 5 Hasil Stopword Removal

Sebelum Stopword Removal	Setelah Stopword Removal
[pengiriman, produk, sangat, cepat, dan, kualitas, barang, memuaskan]	[pengiriman, produk, cepat, kualitas, barang, memuaskan]
[produk, sudah, diterima, sesuai, pesanan]	[produk, diterima, sesuai, pesanan]
[barang, datang, terlambat, dan, kemasan, rusak]	[barang, datang, terlambat, kemasan, rusak]

5) Stemming

Tahapan stemming dilakukan untuk mengubah kata berimbuhan menjadi kata dasar menggunakan algoritma stemming Bahasa Indonesia. Penelitian Rosid et al. (2020) menunjukkan bahwa stemming dapat

meningkatkan kualitas preprocessing pada dokumen teks berbahasa Indonesia(Rosid et al., 2020; Roiqoh et al., 2023).

Tabel 6. Hasil Stemming

Sebelum Stemming	Setelah Stemming
[pengiriman, produk, cepat, kualitas, barang, memuaskan]	[kirim, produk, cepat, kualitas, barang, puas]
[produk, diterima, sesuai, pesanan]	[produk, terima, sesuai, pesan]
[barang, datang, terlambat, kemasan, rusak]	[barang, datang, lambat, kemas, rusak]

Berdasarkan hasil preprocessing, setiap ulasan mengalami transformasi bertahap mulai dari proses cleaning, case folding, tokenizing, stopword removal, hingga stemming. Tahapan tersebut berhasil mengurangi noise pada data teks dan menghasilkan representasi data yang lebih bersih, konsisten, serta terstruktur. Kondisi ini meningkatkan kualitas ekstraksi fitur menggunakan TF-IDF, karena kata-kata yang memiliki makna penting dapat direpresentasikan dengan lebih baik. Dengan demikian, proses klasifikasi menggunakan algoritma Naïve Bayes dan K-Nearest Neighbor (KNN) dapat mengenali pola sentimen secara lebih akurat dan menghasilkan performa klasifikasi yang lebih optimal.

D. Penentuan Label Sentimen

Sebelum proses klasifikasi dilakukan, seluruh data ulasan pelanggan terlebih dahulu diberikan label sentimen ke dalam tiga kategori, yaitu positif, negatif, dan netral. Proses pelabelan dilakukan secara manual berdasarkan isi dan konteks ulasan pelanggan. Ulasan yang mengandung apresiasi, kepuasan, atau pengalaman positif terhadap layanan dikategorikan sebagai sentimen positif. Ulasan yang berisi keluhan, kritik, atau ketidakpuasan dikategorikan sebagai sentimen negatif. Sementara itu, ulasan yang bersifat informatif, deskriptif, atau tidak menunjukkan kecenderungan emosi tertentu dikategorikan sebagai sentimen netral.

Proses pelabelan dilakukan sebagai tahap awal untuk membentuk dataset yang siap digunakan dalam proses pelatihan dan pengujian model klasifikasi. Hasil pelabelan kemudian digunakan sebagai kelas target (target class) pada algoritma Naïve Bayes dan K-Nearest Neighbor (KNN).

E. Pembobotan Fitur Menggunakan TF-IDF

Setelah melalui tahap preprocessing, data teks diubah menjadi representasi numerik menggunakan metode Term Frequency–Inverse Document Frequency (TF-IDF). Metode ini digunakan untuk memberikan bobot pada setiap kata berdasarkan tingkat kemunculannya dalam dokumen.

TF-IDF dihitung menggunakan persamaan (1)

$$TF - IDF(t,d) = TF(t,d) \times IDF(t) \qquad (1)$$

Dimana TF (Term Frequency) menunjukkan frekuensi kemunculan kata pada suatu dokumen sedangkan IDF (Inverse Document Frequency) menunjukkan tingkat kepentingan kata terhadap keseluruhan dokumen.

Selanjutnya, nilai Inverse Document Frequency (IDF) dihitung menggunakan Persamaan (2).

$$IDF\ (t) = \log\left(\frac{N}{df(t)}\right) \qquad (2)$$

Dengan (N) merupakan jumlah seluruh dokumen pada dataset dan $df(t)$ merupakan jumlah dokumen yang mengandung term t.

Metode TF-IDF dipilih karena mampu meningkatkan kualitas representasi teks serta mengurangi pengaruh kata-kata yang terlalu sering muncul tetapi kurang informatif. Dengan demikian, fitur yang memiliki relevansi tinggi terhadap sentimen pelanggan akan memperoleh bobot yang lebih besar sehingga dapat meningkatkan performa proses klasifikasi.

F. Klasifikasi Menggunakan Algoritma Naïve Bayes

Algoritma Naïve Bayes merupakan salah satu metode klasifikasi yang didasarkan pada teori probabilitas dan Teorema Bayes. Algoritma ini bekerja dengan menghitung probabilitas suatu data termasuk ke dalam kelas tertentu berdasarkan informasi dari atribut atau fitur yang dimiliki. Naïve Bayes memiliki asumsi bahwa setiap fitur bersifat independen terhadap fitur lainnya (conditional independence assumption), sehingga proses perhitungan probabilitas menjadi lebih sederhana dan efisien.

Metode ini banyak digunakan dalam analisis sentimen karena mampu menangani data teks dengan jumlah fitur yang besar serta memiliki

waktu komputasi yang relatif cepat. Pada proses klasifikasi, algoritma akan menghitung probabilitas posterior dari setiap kelas sentimen, kemudian memilih kelas dengan nilai probabilitas tertinggi sebagai hasil prediksi.

Persamaan dasar Teorema Bayes yang digunakan dalam algoritma Naïve Bayes adalah sebagai berikut:

$$P(C|X) = \frac{P(X|C)P(C)}{P(X)} \quad (3)$$

Dalam penelitian ini, algoritma Naïve Bayes digunakan untuk mengklasifikasikan ulasan pelanggan PT XYZ ke dalam kategori sentimen positif, negatif, dan netral berdasarkan fitur hasil ekstraksi TF-IDF. Hasil klasifikasi kemudian digunakan untuk mengevaluasi kemampuan algoritma dalam memahami opini pengguna terhadap aplikasi yang dianalisis.

G. Klasifikasi Menggunakan Algoritma K-Nearest Neighbor (KNN)

Algoritma K-Nearest Neighbor (KNN) merupakan metode klasifikasi berbasis instance yang mengelompokkan data baru berdasarkan kedekatan jarak dengan sejumlah data tetangga terdekat pada data latih. Prinsip utama algoritma ini adalah bahwa objek yang memiliki karakteristik serupa cenderung berada pada kelas yang sama. Pada penelitian ini, proses klasifikasi dilakukan dengan menentukan kelas mayoritas dari sejumlah tetangga terdekat yang memiliki jarak paling kecil terhadap data yang akan diklasifikasikan.

Pengukuran jarak antar data dilakukan menggunakan metode Euclidean Distance yang dirumuskan sebagai berikut:

$$d(x,y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \quad (4)$$

Dimana $d(x,y)$ menyatakan jarak antara dua data yang dibandingkan, x_i merupakan nilai atribut ke-i pada data pertama, y_i merupakan nilai atribut ke-i pada data kedua, sedangkann n menunjukkan jumlah atribut atau fitur yang digunakan dalam proses perhitungan jarak. Nilai jarak yang lebih kecil menunjukkan tingkat kemiripan yang lebih tinggi antara kedua data.

Nilai parameter *K* ditentukan melalui proses eksperimen menggunakan beberapa alternatif nilai K untuk memperoleh performa klasifikasi terbaik. Berdasarkan penelitian terdahulu dan hasil eksperimen awal pada penelitian ini, nilai K = 5 memberikan performa klasifikasi terbaik sehingga digunakan sebagai parameter pada penelitian ini. Nilai $K = 5$ dipilih berdasarkan beberapa penelitian terdahulu yang menunjukkan bahwa nilai tersebut mampu memberikan keseimbangan antara kompleksitas model dan tingkat akurasi klasifikasi pada data teks berbahasa Indonesia.

H. Evaluasi Model

Evaluasi model dilakukan menggunakan Confusion Matrix untuk membandingkan label aktual dengan label hasil prediksi.

Confusion Matrix menghasilkan empat komponen utama: True Positive (TP), True Negative (TN), False Positive (FP), False Negative (FN).

Berdasarkan nilai tersebut dihitung metrik evaluasi sebagai berikut.

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

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

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

$$f1 - Score = 2\ x\ \frac{Precision\ x\ Recall}{Precision\ x\ Recall} \quad (8)$$

I. Metode Analisis Data

Metode analisis dilakukan dengan membandingkan performa algoritma Naïve Bayes dan K-Nearest Neighbor berdasarkan nilai accuracy, precision, recall, dan F1-score. Selain itu, dilakukan analisis confusion matrix untuk mengidentifikasi pola kesalahan klasifikasi pada masing-masing algoritma sehingga diperoleh pemahaman yang lebih komprehensif mengenai keunggulan dan keterbatasan kedua model. Algoritma dengan performa terbaik direkomendasikan sebagai model klasifikasi sentimen pada PT XYZ.

III. HASIL DAN PEMBAHASAN

A. Deskripsi Dataset

Penelitian ini menggunakan dataset berupa 3.500 ulasan pelanggan internal PT XYZ yang telah diberi label sentimen ke dalam tiga kategori, yaitu positif, negatif, dan netral. Dataset berisi opini pelanggan mengenai kualitas produk, pelayanan pelanggan, kecepatan pengiriman, harga, kemudahan transaksi, serta pengalaman penggunaan platform secara keseluruhan.

Pada penelitian ini, dataset dibagi menggunakan metode split validation dengan rasio 80:20, dimana sebanyak 2.800 data (80%)

digunakan sebagai data pelatihan (training data) dan 700 data (20%) digunakan sebagai data pengujian (testing data). Pembagian data dilakukan untuk memastikan model dapat diuji secara objektif terhadap data yang belum pernah dipelajari sebelumnya.

Pemilihan metode split validation didasarkan pada jumlah dataset yang relatif besar sehingga mampu menyediakan data pelatihan yang cukup untuk membangun model sekaligus mempertahankan data pengujian yang representatif dalam mengevaluasi kemampuan generalisasi model. Selain itu, metode ini memiliki waktu komputasi yang lebih efisien dibandingkan metode validasi berulang sehingga sesuai untuk penelitian komparatif yang membandingkan performa algoritma Naïve Bayes dan K-Nearest Neighbor (KNN).

Tabel 7 Distribusi Dataset Penelitian

Dataset	Jumlah	Persentase
Training	2.800	80%
Testing	700	20%
Total	3.500	100%

Sumber: Hasil Diolah penulis (2026)

Pembagian dataset dilakukan secara proporsional agar model klasifikasi memiliki kemampuan generalisasi yang baik dalam memprediksi sentimen pelanggan pada data baru.

B. Hasil Text Preprocessing

Tahapan preprocessing dilakukan untuk meningkatkan kualitas data teks sebelum dilakukan proses klasifikasi. Tahapan ini bertujuan menghilangkan noise, menyeragamkan struktur data, serta menghasilkan representasi fitur yang lebih optimal.

Tahapan preprocessing yang diterapkan dalam penelitian ini meliputi case folding, cleaning, tokenizing, stopword removal, dan stemming.

Tabel 8 Hasil Preprocessing

Tahapan	Hasil
Data Asli	“Pengiriman barang sangat cepat, kualitas produk bagus banget!!!”
Cleaning	“pengiriman barang sangat cepat kualitas produk bagus banget”
Case Folding	“pengiriman barang sangat cepat, kualitas produk bagus banget!!!”
Tokenization	[pengiriman, barang, sangat, cepat, kualitas, produk, bagus, banget]
Stopword Removal	[pengiriman, barang, cepat, kualitas, produk, bagus]
Stemming	[kirim, barang, cepat, kualitas, produk, bagus]

Berdasarkan Tabel 8, proses preprocessing berhasil membersihkan data dari karakter yang tidak relevan seperti tanda baca dan kata-kata umum (stopwords) sehingga menghasilkan data teks yang lebih terstruktur. Tahapan stemming juga membantu menyederhanakan kata berimbuhan menjadi bentuk dasar sehingga mampu meningkatkan konsistensi representasi fitur pada proses klasifikasi

C. Hasil Pembobotan Kata Menggunakan TF-IDF

Setelah proses preprocessing, data teks ditransformasikan ke bentuk numerik menggunakan metode Term Frequency–Inverse Document Frequency (TF-IDF). Metode ini digunakan untuk menentukan bobot kepentingan setiap kata berdasarkan frekuensi kemunculannya pada dokumen.

Tabel 9 Hasil Pembobotan TF-IDF

Term	TF	IDF	TF-IDF
cepat	7	1.32	9.24
bagus	9	1.25	11.25
lambat	4	1.78	7.12
mahal	5	1.51	7.55
puas	8	1.43	11.44

Hasil pembobotan TF-IDF menunjukkan bahwa kata-kata yang memiliki relevansi tinggi terhadap opini pelanggan memperoleh bobot yang lebih besar sehingga dapat meningkatkan akurasi model klasifikasi sentimen. Pembobotan TF-IDF menghasilkan representasi numerik yang mampu membedakan tingkat kepentingan setiap kata dalam dokumen. Kata yang sering muncul pada suatu ulasan tetapi jarang ditemukan pada keseluruhan dataset akan memperoleh bobot lebih tinggi sehingga memberikan kontribusi yang lebih besar dalam proses klasifikasi. Sebaliknya, kata-kata umum yang muncul pada hampir seluruh dokumen akan memiliki bobot lebih rendah karena dianggap kurang mampu membedakan kategori sentimen. Dengan

representasi fitur yang lebih informatif, algoritma Naïve Bayes maupun K-Nearest Neighbor (KNN) dapat mengenali pola sentimen secara lebih efektif sehingga menghasilkan proses klasifikasi yang lebih akurat.

D. Hasil Pengujian Algoritma Naïve Bayes

Pengujian algoritma Naïve Bayes dilakukan menggunakan 700 data ulasan pelanggan yang telah melalui tahapan preprocessing meliputi case folding, tokenizing, stopword removal, dan stemming. Selanjutnya, data teks direpresentasikan menggunakan metode Term Frequency-Inverse Document Frequency (TF-IDF) untuk menghasilkan bobot fitur yang digunakan dalam proses klasifikasi sentimen.

Hasil evaluasi performa algoritma Naïve Bayes ditunjukkan pada Tabel 10.

Tabel 10 Hasil Evaluasi Naïve Bayes

Metrik Evaluasi	Nilai
Accuracy	86.85%
Precision	86.30%
Recall	88.50%
F1-Score	87.30%

Berdasarkan Tabel 10, algoritma Naïve Bayes memperoleh tingkat accuracy sebesar 86,85%, yang menunjukkan bahwa model mampu mengklasifikasikan sebagian besar data ulasan pelanggan dengan baik. Nilai precision sebesar 86,30% menunjukkan bahwa sebagian besar prediksi yang dihasilkan sesuai dengan kelas sebenarnya, sedangkan nilai recall sebesar 88,50% mengindikasikan bahwa model mampu mengidentifikasi sebagian besar data sentimen yang terdapat pada dataset. Sementara itu, F1-Score sebesar 87,30% menunjukkan keseimbangan yang baik antara nilai precision dan recall, sehingga model memiliki performa klasifikasi yang cukup baik.

Hasil tersebut menunjukkan bahwa algoritma Naïve Bayes mampu mengklasifikasikan sentimen secara efektif melalui pendekatan probabilistik yang menghitung peluang kemunculan setiap fitur terhadap masing-masing kelas sentimen. Meskipun demikian, algoritma ini mengasumsikan bahwa setiap fitur bersifat independen, sedangkan pada data teks hubungan antar kata sering kali saling memengaruhi dalam membentuk konteks suatu kalimat. Kondisi tersebut menyebabkan asumsi independensi tidak selalu terpenuhi sehingga dapat memengaruhi akurasi klasifikasi, terutama pada ulasan yang mengandung kalimat kompleks atau memiliki sentimen campuran. Namun demikian, Naïve Bayes tetap memberikan performa yang kompetitif karena memiliki proses komputasi yang sederhana, efisien, dan mampu menghasilkan klasifikasi dengan waktu pemrosesan yang relatif cepat.

Tebel 11 Confusion Matrix Naïve Bayes

Actual / Predicted	Positif	Negatif	Netral
Positif	238	17	11
Negatif	19	177	13
Netral	15	17	193

Berdasarkan Tabel 11, sebagian besar data berhasil diklasifikasikan dengan benar pada setiap kategori sentimen. Namun, masih terdapat beberapa kesalahan klasifikasi, terutama pada ulasan yang mengandung kata-kata ambigu atau kombinasi sentimen positif dan negatif dalam satu kalimat. Kondisi ini menunjukkan bahwa Naïve Bayes masih memiliki keterbatasan dalam memahami konteks dan hubungan semantik antar kata karena proses klasifikasi didasarkan pada probabilitas kemunculan setiap fitur secara independen. Akibatnya, beberapa ulasan netral diprediksi sebagai positif maupun negatif ketika mengandung kata-kata evaluatif yang memiliki kecenderungan sentimen tertentu. Temuan ini menunjukkan bahwa meskipun Naïve Bayes memiliki efisiensi komputasi yang tinggi, kemampuan model dalam menangani konteks bahasa yang kompleks masih perlu ditingkatkan.

E. Hasil Pengujian Algoritma K-Nearest Neighbor (KNN)

Pengujian algoritma K-Nearest Neighbor (KNN) dilakukan menggunakan dataset yang sama dengan algoritma Naïve Bayes untuk memastikan proses evaluasi berjalan secara objektif. Pada penelitian ini digunakan nilai parameter $K = 5$, dimana proses klasifikasi dilakukan berdasarkan mayoritas kelas dari lima tetangga terdekat yang memiliki kemiripan tertinggi terhadap data uji.

Berdasarkan beberapa penelitian terdahulu, nilai $K = 5$ sering digunakan karena mampu memberikan keseimbangan yang baik antara akurasi dan kemampuan generalisasi model. Oleh karena itu, penelitian ini menggunakan nilai $K = 5$ sebagai parameter klasifikasi.

Hasil evaluasi performa algoritma KNN disajikan pada Tabel 12.

Tabel 12 Hasil Evaluasi Algoritma KNN

Metrik	Nilai
Accuracy	92.57%
Precision	91.56%
Recall	92.73%
F1-Score	92.14%

Berdasarkan Tabel 12, algoritma K-Nearest Neighbor (KNN) memperoleh nilai accuracy sebesar 92,57%, yang lebih tinggi dibandingkan algoritma Naïve Bayes. Nilai precision sebesar 91,56% menunjukkan bahwa sebagian besar prediksi yang dihasilkan model sesuai dengan kelas sebenarnya, sedangkan nilai recall sebesar 92,73% mengindikasikan kemampuan model yang sangat baik dalam mengidentifikasi seluruh data sentimen pelanggan. Selain itu, nilai F1-Score sebesar 92,14% menunjukkan keseimbangan yang baik antara precision dan recall, sehingga model memiliki performa klasifikasi yang sangat baik.

Hasil tersebut menunjukkan bahwa algoritma K-Nearest Neighbor (KNN) memiliki performa yang lebih tinggi dibandingkan Naïve Bayes. Hal ini menunjukkan bahwa pendekatan berbasis kemiripan (similarity-based classification) lebih sesuai untuk karakteristik data ulasan pelanggan yang memiliki variasi kosakata dan pola kalimat yang beragam. Dengan memanfaatkan kedekatan antar dokumen berdasarkan representasi fitur TF-IDF, KNN mampu mengidentifikasi pola sentimen secara lebih akurat tanpa mengasumsikan independensi antar fitur. Representasi fitur menggunakan TF-IDF juga membantu meningkatkan kualitas pengukuran jarak antar dokumen sehingga proses pencarian tetangga terdekat menjadi lebih optimal. Oleh karena itu, model KNN mampu menghasilkan klasifikasi yang lebih konsisten dan akurat dibandingkan pendekatan probabilistik yang digunakan pada algoritma Naïve Bayes.

Tabel 13 Confusion Matrix KNN

Aktual/Prediksi	Positif	Netral	Negatif
Positif	251	8	7
Negatif	10	188	11
Netral	6	10	209

Berdasarkan Tabel 13, jumlah prediksi benar yang dihasilkan algoritma KNN lebih tinggi dibandingkan Naïve Bayes pada seluruh kategori sentimen. Kelas positif berhasil diklasifikasikan dengan benar sebanyak 251 data, kelas negatif sebanyak 188 data, dan kelas netral sebanyak 209 data, sedangkan jumlah kesalahan klasifikasi relatif lebih sedikit pada setiap kategori. Hasil ini menunjukkan bahwa pendekatan berbasis kemiripan yang digunakan KNN lebih efektif dalam mengenali karakteristik ulasan pelanggan, terutama ketika terdapat variasi penggunaan kata dan konteks kalimat. Kemampuan tersebut memungkinkan model menghasilkan prediksi yang lebih akurat dan memiliki kemampuan generalisasi yang lebih baik terhadap data uji dibandingkan algoritma Naïve Bayes.

F. Perbandingan Performa Algoritma

Tabel 14 Perbandingan Kinerja Model

Algoritma	Accuracy	Precison	Recall	F1-Score
Naïve Bayes	86.85%	86.30%	88.50%	87.30%
KNN	92.57%	91.56%	92.73%	92.14%

Berdasarkan Tabel 14, algoritma K-Nearest Neighbor (KNN) menunjukkan performa yang lebih unggul dibandingkan Naïve Bayes pada seluruh metrik evaluasi. KNN memperoleh nilai accuracy sebesar 92,57%, lebih tinggi dibandingkan Naïve Bayes yang memperoleh 86,85%, sehingga terdapat selisih akurasi sebesar 5,72%. Selain itu, nilai precision, recall, dan F1-score pada KNN juga lebih tinggi masing-masing sebesar 91,56%, 92,73%, dan 92,14%, sedangkan Naïve Bayes memperoleh 86,30%, 88,50%, dan 87,30%. Hasil tersebut menunjukkan bahwa algoritma KNN memiliki kemampuan yang lebih baik dalam mengidentifikasi dan mengklasifikasikan sentimen pelanggan dibandingkan algoritma Naïve Bayes pada dataset yang digunakan dalam penelitian ini.

Perbedaan performa tersebut mengindikasikan bahwa algoritma K-Nearest Neighbor lebih efektif dalam mengenali pola sentimen pada dataset penelitian ini. Hal ini disebabkan KNN melakukan proses klasifikasi berdasarkan tingkat kemiripan antar dokumen sehingga mampu memanfaatkan representasi fitur hasil pembobotan TF-IDF secara lebih optimal. Pendekatan berbasis kemiripan (similarity-based classification) memungkinkan KNN mempertahankan karakteristik lokal setiap dokumen sehingga hubungan antar kata yang membentuk konteks kalimat tetap dapat direpresentasikan dengan baik. Sebaliknya, Naïve Bayes mengandalkan asumsi independensi

antar fitur yang kurang sesuai dengan karakteristik data teks, karena hubungan antar kata sering kali saling memengaruhi dalam membentuk makna suatu kalimat.Kondisi tersebut diduga menjadi salah satu faktor yang menyebabkan performa Naïve Bayes lebih rendah dibandingkan KNN pada penelitian ini.

Hasil penelitian ini sejalan dengan penelitian Elfansyah et al. (2024) yang membandingkan algoritma K-Nearest Neighbor (KNN) dan Naïve Bayes pada analisis sentimen pengguna aplikasi DANA menggunakan pembobotan TF-IDF(Rayhan et al., 2024). Penelitian tersebut menunjukkan bahwa algoritma KNN memberikan performa yang lebih baik dibandingkan Naïve Bayes dalam mengklasifikasikan sentimen pengguna. Temuan serupa juga dilaporkan oleh Pratmanto et al. (2023) pada analisis sentimen ulasan aplikasi Identitas Kependudukan Digital (IKD), di mana KNN menghasilkan tingkat akurasi dan presisi yang lebih tinggi dibandingkan Naïve Bayes(Pratmanto et al., 2023). Temuan tersebut menunjukkan bahwa hasil penelitian ini konsisten dengan penelitian terdahulu yang melaporkan bahwa algoritma KNN mampu memberikan performa klasifikasi yang lebih baik dibandingkan Naïve Bayes pada data analisis sentimen berbasis TF-IDF.

Berdasarkan hasil pengujian pada dataset penelitian ini, algoritma KNN merupakan metode yang lebih sesuai dibandingkan Naïve Bayes untuk melakukan klasifikasi sentimen ulasan pelanggan PT XYZ. Dari sisi implementasi, penggunaan algoritma KNN berpotensi meningkatkan ketepatan identifikasi opini pelanggan apabila diterapkan pada karakteristik data yang serupa. Informasi sentimen yang dihasilkan dapat dimanfaatkan sebagai dasar evaluasi kualitas layanan, penyusunan strategi peningkatan kepuasan pelanggan, serta mendukung proses pengambilan keputusan berbasis data (data-driven decision making) di lingkungan e-commerce.

G. Pembahasan

Hasil penelitian menunjukkan bahwa algoritma K-Nearest Neighbor (KNN) memberikan performa klasifikasi sentimen yang lebih baik dibandingkan Naïve Bayes pada dataset ulasan pelanggan PT XYZ. Perbedaan performa tersebut menunjukkan bahwa efektivitas suatu algoritma klasifikasi tidak hanya ditentukan oleh kompleksitas metode yang digunakan, tetapi juga sangat dipengaruhi oleh karakteristik dataset, representasi fitur, serta hubungan antar kata dalam dokumen teks. Pada penelitian ini, karakteristik ulasan pelanggan yang mengandung variasi kosakata, ekspresi, dan konteks bahasa menyebabkan pendekatan berbasis kemiripan (similarity-based classification) yang digunakan oleh KNN lebih mampu mengenali pola sentimen dibandingkan pendekatan probabilistik yang diterapkan pada Naïve Bayes.

Secara teoritis, algoritma Naïve Bayes bekerja berdasarkan Teorema Bayes dengan asumsi bahwa setiap fitur bersifat independen (conditional independence assumption). Asumsi tersebut menyederhanakan proses perhitungan probabilitas sehingga algoritma memiliki keunggulan dalam kecepatan komputasi dan efisiensi ketika menangani data berdimensi tinggi. Namun, pada data teks, hubungan antar kata umumnya tidak bersifat independen karena makna suatu kalimat dibentuk oleh keterkaitan antar kata. Akibatnya, asumsi independensi tersebut tidak selalu terpenuhi sehingga dapat mengurangi kemampuan Naïve Bayes dalam memahami konteks kalimat, terutama pada ulasan yang mengandung frasa, negasi, atau kombinasi sentimen positif dan negatif dalam satu dokumen.

Sebaliknya, algoritma K-Nearest Neighbor (KNN) melakukan klasifikasi berdasarkan tingkat kemiripan antar dokumen tanpa mengasumsikan adanya independensi antar fitur. Pada penelitian ini, setiap dokumen direpresentasikan dalam bentuk vektor menggunakan metode Term Frequency–Inverse Document Frequency (TF-IDF) sehingga setiap kata memperoleh bobot sesuai tingkat kepentingannya dalam keseluruhan koleksi dokumen. Representasi tersebut menghasilkan ruang vektor yang mampu menggambarkan karakteristik setiap ulasan secara lebih informatif. Melalui perhitungan jarak antar dokumen pada ruang vektor tersebut, KNN dapat mengidentifikasi ulasan yang memiliki pola kata serupa sehingga proses penentuan kelas sentimen menjadi lebih akurat. Dengan demikian, kombinasi TF-IDF dan KNN mampu memanfaatkan karakteristik lokal setiap dokumen secara optimal dalam proses klasifikasi.

Penggunaan TF-IDF juga memberikan kontribusi penting terhadap peningkatan kualitas klasifikasi. Pembobotan TF-IDF mengurangi pengaruh kata-kata umum yang sering muncul tetapi kurang memiliki kemampuan membedakan

kategori sentimen, sekaligus meningkatkan bobot kata-kata yang lebih spesifik terhadap suatu opini pelanggan. Representasi fitur yang lebih diskriminatif tersebut memungkinkan algoritma klasifikasi memperoleh informasi yang lebih relevan selama proses pembelajaran maupun prediksi. Karakteristik tersebut memberikan keuntungan yang lebih besar bagi KNN karena proses pencarian tetangga terdekat sangat bergantung pada kualitas representasi fitur dan pengukuran kemiripan antar dokumen.

Temuan penelitian ini sejalan dengan penelitian Rayhan et al. (2024) yang melaporkan bahwa kombinasi algoritma K-Nearest Neighbor dan TF-IDF mampu menghasilkan performa klasifikasi yang baik pada analisis sentimen ulasan pengguna aplikasi DANA (Rayhan et al., 2024). Hasil penelitian ini juga mendukung temuan Muslim et al. (2024) yang menunjukkan bahwa KNN memberikan performa yang lebih baik dibandingkan Naïve Bayes pada analisis sentimen ulasan pengguna aplikasi CapCut (Muslim et al., 2024). Kesamaan hasil tersebut menunjukkan bahwa pendekatan berbasis kemiripan cenderung lebih efektif ketika diterapkan pada data teks yang memiliki variasi kosakata dan pola kalimat yang tinggi.

Namun demikian, hasil penelitian ini berbeda dengan penelitian Dany Pratmanto yang menemukan bahwa Naïve Bayes memiliki performa lebih baik dibandingkan KNN pada analisis sentimen aplikasi Vidio. Perbedaan hasil tersebut menunjukkan bahwa efektivitas algoritma klasifikasi sangat dipengaruhi oleh karakteristik dataset, distribusi kelas, jumlah data, teknik preprocessing, serta metode representasi fitur yang digunakan.

Di sisi lain, hasil penelitian ini berbeda dengan penelitian Pratmanto et al. (2024) yang melaporkan bahwa Naïve Bayes memberikan performa yang lebih baik dibandingkan KNN pada analisis sentimen ulasan aplikasi Vidio(Pratmanto et al., 2024). Perbedaan tersebut menunjukkan bahwa tidak terdapat satu algoritma yang selalu memberikan hasil terbaik pada seluruh kasus analisis sentimen. Performa algoritma sangat dipengaruhi oleh karakteristik dataset, distribusi kelas, jumlah data, proses text preprocessing, teknik representasi fitur, serta parameter yang digunakan selama proses klasifikasi. Dengan demikian, pemilihan algoritma klasifikasi sebaiknya disesuaikan dengan karakteristik data yang akan dianalisis, bukan hanya berdasarkan hasil penelitian sebelumnya.

Berdasarkan keseluruhan hasil penelitian, dapat dipahami bahwa kombinasi text preprocessing, pembobotan TF-IDF, dan algoritma K-Nearest Neighbor mampu menghasilkan model klasifikasi sentimen yang lebih sesuai untuk karakteristik dataset ulasan pelanggan PT XYZ. Temuan ini memberikan bukti empiris bahwa pendekatan berbasis kemiripan dapat dimanfaatkan sebagai alternatif yang efektif dalam pengembangan sistem analisis sentimen pada perusahaan e-commerce. Informasi sentimen yang dihasilkan dapat mendukung proses evaluasi kualitas layanan, identifikasi kebutuhan pelanggan, serta pengambilan keputusan berbasis data (data-driven decision making) untuk meningkatkan kepuasan pelanggan.

H. Implikasi Penelitian

Hasil penelitian ini memberikan implikasi teoritis dan praktis dalam pengembangan analisis sentimen berbasis machine learning pada sektor e-commerce. Secara teoritis, penelitian ini memperkuat bukti empiris bahwa algoritma K-Nearest Neighbor (KNN) mampu memberikan performa klasifikasi yang lebih baik dibandingkan Naïve Bayes pada data teks yang direpresentasikan menggunakan metode TF-IDF. Temuan ini menunjukkan bahwa pemilihan algoritma klasifikasi berpengaruh terhadap kualitas hasil analisis sentimen, sehingga dapat menjadi referensi bagi penelitian selanjutnya yang mengembangkan maupun membandingkan berbagai algoritma machine learning untuk klasifikasi data teks.

Selain itu, penelitian ini memberikan kontribusi ilmiah dengan menunjukkan bahwa kombinasi TF-IDF dan KNN menghasilkan performa klasifikasi yang lebih baik dibandingkan kombinasi TF-IDF dan Naïve Bayes pada dataset ulasan pelanggan PT XYZ. Temuan ini memperkaya kajian mengenai penerapan algoritma klasifikasi pada analisis sentimen serta dapat menjadi dasar bagi penelitian selanjutnya untuk mengeksplorasi algoritma lain, seperti Support Vector Machine (SVM), Random Forest, Decision Tree, maupun metode berbasis Deep Learning.

Secara praktis, hasil penelitian menunjukkan bahwa algoritma K-Nearest Neighbor layak dipertimbangkan sebagai model utama dalam pengembangan sistem analisis sentimen

pelanggan di PT XYZ. Dengan tingkat accuracy sebesar 92,57%, implementasi algoritma ini berpotensi membantu perusahaan dalam mengidentifikasi opini pelanggan secara lebih cepat dan akurat dibandingkan proses analisis manual. Informasi sentimen yang dihasilkan dapat dimanfaatkan sebagai dasar evaluasi kualitas produk dan layanan, penyusunan strategi peningkatan kepuasan pelanggan, serta mendukung proses pengambilan keputusan berbasis data (data-driven decision making).

Selain itu, sistem analisis sentimen yang dikembangkan memungkinkan perusahaan mendeteksi keluhan pelanggan secara lebih dini sehingga tindakan perbaikan dapat dilakukan secara lebih cepat dan tepat sasaran. Dengan demikian, implementasi algoritma KNN tidak hanya mendukung peningkatan kualitas layanan, tetapi juga berpotensi memperkuat loyalitas pelanggan dan meningkatkan daya saing perusahaan di tengah persaingan industri e-commerce yang semakin kompetitif.

I. Keterbatasan Penelitian

Penelitian ini menggunakan dataset sebanyak 3.500 ulasan pelanggan dengan pembagian 2.800 data latih dan 700 data uji menggunakan metode Split Validation dengan rasio 80:20. Meskipun jumlah data tersebut mampu menghasilkan performa klasifikasi yang baik, penggunaan dataset yang lebih besar dan lebih beragam berpotensi meningkatkan kemampuan generalisasi model terhadap data baru. Selain itu, evaluasi model pada penelitian ini hanya dilakukan menggunakan satu skenario pembagian data sehingga hasil yang diperoleh masih berpotensi dipengaruhi oleh komposisi data latih dan data uji. Oleh karena itu, penelitian selanjutnya disarankan menerapkan metode k-fold cross-validation untuk memperoleh evaluasi performa model yang lebih stabil dan komprehensif.

Selain itu, penelitian ini hanya membandingkan dua algoritma klasifikasi, yaitu Naïve Bayes dan K-Nearest Neighbor (KNN). Meskipun KNN menunjukkan performa terbaik dengan nilai accuracy sebesar 92,57%, algoritma ini memiliki keterbatasan berupa meningkatnya waktu komputasi seiring bertambahnya jumlah data karena seluruh data pelatihan digunakan pada saat proses klasifikasi. Di sisi lain, penelitian ini masih menggunakan satu teknik pembobotan fitur, yaitu Term Frequency–Inverse Document Frequency (TF-IDF). Oleh karena itu, penelitian selanjutnya dapat mengeksplorasi metode representasi fitur lain, seperti Word2Vec, FastText, BERT, atau model berbasis Transformer, serta membandingkannya dengan algoritma lain seperti Support Vector Machine (SVM), Random Forest, maupun metode Deep Learning untuk memperoleh performa klasifikasi yang lebih optimal.

IV. KESIMPULAN

Berdasarkan hasil penelitian, dapat disimpulkan bahwa penerapan teknik text preprocessing dan pembobotan Term Frequency–Inverse Document Frequency (TF-IDF) mampu menghasilkan representasi fitur yang efektif untuk proses klasifikasi sentimen pelanggan e-commerce PT XYZ. Kedua algoritma yang diuji, yaitu Naïve Bayes dan K-Nearest Neighbor (KNN), menunjukkan kemampuan yang baik dalam mengklasifikasikan sentimen pelanggan ke dalam kategori positif, negatif, dan netral.

Hasil evaluasi menunjukkan bahwa algoritma K-Nearest Neighbor (KNN) dengan nilai K = 5 memberikan performa terbaik dengan nilai accuracy sebesar 92,57%, precision 91,56%, recall 92,73%, dan F1-score 92,14%. Sementara itu, algoritma Naïve Bayes memperoleh accuracy sebesar 86,85%, precision 86,30%, recall 88,50%, dan F1-score 87,30%. Dengan demikian, KNN menghasilkan tingkat accuracy yang lebih tinggi sebesar 5,72% dibandingkan Naïve Bayes. Hasil tersebut menunjukkan bahwa pendekatan berbasis kemiripan (similarity-based classification) lebih efektif dalam mengenali pola sentimen pada data ulasan pelanggan PT XYZ dibandingkan pendekatan probabilistik yang digunakan oleh Naïve Bayes.

Kontribusi utama penelitian ini adalah memberikan bukti empiris melalui evaluasi komparatif algoritma Naïve Bayes dan K-Nearest Neighbor (KNN) menggunakan dataset ulasan pelanggan internal PT XYZ dengan pembobotan TF-IDF. Hasil penelitian menunjukkan bahwa algoritma KNN memberikan performa klasifikasi yang lebih baik dibandingkan Naïve Bayes sehingga layak dipertimbangkan sebagai metode klasifikasi sentimen pada lingkungan e-commerce.

Untuk penelitian selanjutnya, disarankan menggunakan dataset yang lebih besar dan beragam, menerapkan metode validasi yang lebih komprehensif seperti k-fold cross-validation, serta membandingkan performa KNN dan Naïve

Bayes dengan algoritma lain, seperti Support Vector Machine (SVM), Random Forest, maupun metode berbasis Deep Learning, sehingga dapat diperoleh model klasifikasi sentimen dengan tingkat akurasi dan kemampuan generalisasi yang lebih baik.

DAFTAR PUSTAKA

Abdillah, T., Khaira, U., & Hutabarat, B. F. (2024). Komparasi Metode Naive Bayes dan K-Nearest Neighbors Terhadap Analisis Sentimen Pengguna Aplikasi Zenius. *Jurnal PROCESSOR*, *19*(1). https://doi.org/10.33998/processor.2024.19.1.1596

Berlianti, P. M., & Hidayat, E. Y. (2024). Implementasi Naïve Bayes Classifier untuk Sentimen Produk Kecantikan Berdasarkan Ulasan Female Daily. *The Indonesian Journal of Computer Science*, *13*(6), 10248–10256. https://doi.org/10.33022/ijcs.v13i6.4499

Dikananda, F., Rinaldi Dikananda, A., & Anwar, S. (2024). Analisis Sentimen Terhadap Marketplace Menggunakan Algoritma Naïve Bayes. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *8*(4), 8219–8225. https://doi.org/10.36040/jati.v8i4.10965

Fahrezi, A., & Solichin, A. (2024). *Analisis Sentimen Ulasan Aplikasi Mobile Jkn Pada Play Sentiment Analysis of the Jkn Mobile Application Reviews on the Play Store Using the Multinomial Naïve*. *3*(September), 1–10.

Hasbi, B. I., & Putro, I. S. (2025). Perbandingan Kinerja Algoritma Naive Bayes Dan Support Vector Machine untuk Analisis Sentimen Ulasan Produk E- Commerce Menggunakan Pembobotan TF-IDF. *Jurnal Komputer Dan Informatika*, *9*(2), 123–133. https://journals.upi-yai.ac.id/index.php/ikraith-informatika/article/view/4949

Hasibuan, H., Andrianto, R., Budiarti, P., Putri, R., Informasi, S., Teknologi, I., & Lawas, P. (2024). Analisis Sentimen Kepuasan Konsumen E-Commerce Tiktok Shop Menggunakan Metode Naive Bayes. *Jurnal Pendidikan Tambusai*, *8*(2), 33941–33950. https://jptam.org/index.php/jptam/article/view/18809

Latifurrohman, L., Tamrin, T., & Kusumodestoni, R. H. (2025). Perbandingan Metode K-Nearest Neighbor (Knn) Dan Naive Bayes Untuk Analisis Sentimen Review Objek Wisata Jepara Ourland Park (Jop) Berdasarkan Ulasan Pada Google Maps Comparison Of K-Nearest Neighbor (Knn) And Naive Bayes Methods For Sentiment Analy. *JURNAL TEKNIK INFORMATIKA*, *4*(2), 277–286.

Muslim, S. N. S., Nurdiyansyah, F., & Rahman, A. Y. (2024). Perbandingan Algoritma Naive Bayes Dan Knn Dalam Analisis Sentimen Ulasan Pengguna Aplikasi Capcut. *Jurnal Informatika Dan Teknik Elektro Terapan*, *12*(3S1), 3588–3596. https://doi.org/10.23960/jitet.v12i3S1.5156

Ni'mah, A. T., & Arifin, A. Z. (2020). Perbandingan Metode Term Weighting terhadap Hasil Klasifikasi Teks pada Dataset Terjemahan Kitab Hadis. *Rekayasa*, *13*(2), 172–180. https://doi.org/10.21107/rekayasa.v13i2.6412

Nopiah, R., Ekaputri, R. A., Barika, B., & Febriani, R. E. (2024). Impact Of E-Commerce On Indonesia Economic Growth: Intermediation Models With Financial Technology Constraint. *Jurnal REP (Riset Ekonomi Pembangunan)*, *9*(1), 1–23. https://doi.org/10.31002/rep.v9i1.1216

Pratmanto, D., Imaniawan, F. F. D., & Maarif, V. (2023). Analisis Sentimen Pada Ulasan Pengguna Aplikasi Identitas Kependudukan Digital Dengan Metode Naive Bayes Dan K-Nearest. *Computatio : Journal of Computer Science and Information Systems*, *7*(2), 155–166. https://doi.org/10.24912/computatio.v7i2.26322

Pratmanto, D., Widayanto, A., Meisella Kristania, Y., & Wijianto, R. (2024). Analisis Perbandingan Algoritma Naive Bayes Dan KNN Untuk Analisis Sentimen Ulasan Pengguna Aplikasi Vidio Di Google Play Store. *CONTEN : Computer and Network Technology*, *4*(2), 119–124. http://jurnal.bsi.ac.id/index.php/conten

Rayhan, M. R. E., Rudiman, R., & Fendy, F. Y. (2024). Perbandingan Metode K–Nearest Neighbor (Knn) Dan Naive Bayes Terhadap Analisis Sentimen Pada Pengguna E-Wallet

Aplikasi Dana Menggunakan Fitur Ekstraksi TF-IDF. *Jurnal Teknologi Informasi: Jurnal Keilmuan Dan Aplikasi Bidang Teknik Informatika*, *18*(2), 139–159. https://doi.org/10.47111/jti.v18i2.15009

Riki, Eliesse Dameria, S., Hermawan, A., Junaedi, & Kurnia, Y. (2025). Optimizing Sentiment Classification of E-Commerce Product Reviews: A Comparative Study of Naïve Bayes and SVM with SMO. *JUITA: Jurnal Informatika*, *13*(3), 287–295. https://doi.org/10.30595/juita.v13i3.26642

Roiqoh, S., Zaman, B., & Kartono, K. (2023). Analisis Sentimen Berbasis Aspek Ulasan Aplikasi Mobile JKN dengan Lexicon Based dan Naïve Bayes. *Jurnal Media Informatika Budidarma*, *7*(3), 1582. https://doi.org/10.30865/mib.v7i3.6194

Rosid, M. A., Fitrani, A. S., Astutik, I. R. I., Mulloh, N. I., & Gozali, H. A. (2020). Improving Text Preprocessing For Student Complaint Document Classification Using Sastrawi. *IOP Conference Series: Materials Science and Engineering*, *874*(1), 012017. https://doi.org/10.1088/1757-899X/874/1/012017

Situmorang, V. M. E., & Prima, S. R. (2025). The Impact Of E-Commerce On Gross Regional Domestic Product In Indonesia In 2013-2024. *Economous: Journal of Regional Economic Development*, *1*(1), 22–31. https://doi.org/10.33830/economous.v1i1.11429

Wiguna, D. P., & Ginting, S. E. (2026). Analisis Sentimen E-Commerce di Indonesia dengan Algoritma Naive Bayes (Studi Kasus Pada Platform Shopee, Tokopedia, Bukalapak, dan Lazada). *AICOM: Artificial Intelligence and Computing*, *01*(04), 109–114.

Zahra, N. Z., Farhanatussaidah, S., Afifah, N. N., Muthoharoh, L., Satria, A., & Manullang, M. C. T. (2026). Benchmarking Logistic Regression, SVM, Naive Bayes, and IndoBERT Fine-Tuning for Sentiment Analysis on Indonesian Product Reviews. *ArXiv Preprint ArXiv:2605.03439*.

