

Perbandingan Kinerja Algoritma *Naïve Bayes* dan *K-Nearest Neighbors* untuk Analisis Sentimen Ulasan Aplikasi Lazada Menggunakan *Python*

Ahmad Apip^{1*}, Aa.Kurniawan²

¹⁻² Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Pamulang, Indonesia

**Penulis Korespondensi:* ahmadapip1200@gmail.com

Abstract. The rapid growth of e-commerce in Indonesia has led to an increasing number of user reviews shared across various digital platforms, including the Lazada application. These textual reviews contain valuable insights into user satisfaction and experiences but have not been fully utilized for automated sentiment analysis. This study aims to compare the performance of the Naïve Bayes and K-Nearest Neighbor (KNN) algorithms in classifying user sentiment from Lazada reviews collected from the Google Play Store. The data preprocessing stages include text cleansing and case folding, tokenization, stopword removal, and stemming using the Sastrawi library. The cleaned text data were then transformed into numerical representations using the Term Frequency–Inverse Document Frequency (TF-IDF) method before classification. Model performance was evaluated using 10-Fold Cross Validation based on four key metrics: accuracy, precision, recall, and F1-score. The experimental results indicate that the Naïve Bayes algorithm achieved superior performance with an accuracy of 89.56%, precision of 89.53%, recall of 89.56%, and an F1-score of 89.53%. In contrast, the K-Nearest Neighbor (KNN) algorithm obtained an accuracy of 73.22%, precision of 75.21%, recall of 73.22%, and an F1-score of 65.64%. These findings suggest that Naïve Bayes demonstrates higher effectiveness and stability in classifying user sentiment on Lazada reviews compared to the KNN algorithm.

Keywords: *K-Nearest Neighbor; Lazada; Naïve Bayes; Sentiment Analysis; TF-IDF.*

Abstrak. Perkembangan e-commerce di Indonesia mendorong peningkatan jumlah ulasan pengguna yang diunggah melalui berbagai platform digital, termasuk aplikasi Lazada. Ulasan berbentuk teks tersebut mengandung informasi berharga mengenai tingkat kepuasan dan pengalaman pengguna, namun belum dimanfaatkan secara optimal untuk analisis sentimen secara otomatis. Penelitian ini bertujuan untuk membandingkan kinerja algoritma Naïve Bayes dan K-Nearest Neighbor (KNN) dalam mengklasifikasikan sentimen ulasan pengguna aplikasi Lazada berdasarkan data yang diperoleh dari Google Play Store. Tahapan pra-pemrosesan data meliputi pembersihan teks (cleansing dan case folding), tokenisasi, penghapusan stopword (stopword removal), serta stemming menggunakan pustaka Sastrawi. Data teks kemudian direpresentasikan ke dalam bentuk numerik menggunakan metode Term Frequency–Inverse Document Frequency (TF-IDF) sebelum dilakukan proses klasifikasi. Evaluasi performa model dilakukan menggunakan metode 10-Fold Cross Validation dengan empat metrik utama, yaitu akurasi, presisi, recall, dan F1-score. Hasil penelitian menunjukkan bahwa algoritma Naïve Bayes memiliki kinerja lebih unggul dengan akurasi 89,56%, presisi 89,53%, recall 89,56%, dan F1-score 89,53%. Sementara itu, algoritma KNN menghasilkan akurasi 73,22%, presisi 75,21%, recall 73,22%, dan F1-score 65,64%. Berdasarkan hasil tersebut, dapat disimpulkan bahwa Naïve Bayes lebih efektif dan stabil dalam mengklasifikasikan sentimen ulasan pengguna aplikasi Lazada dibandingkan dengan algoritma KNN.

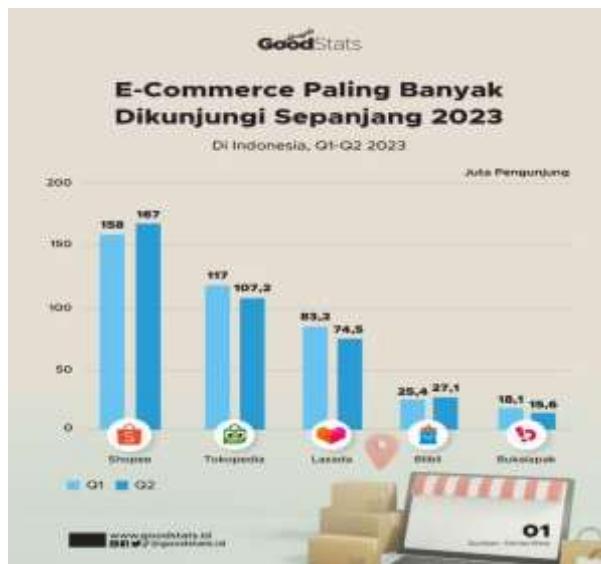
Kata kunci: Analisis Sentimen; K-Nearest Neighbor; Lazada; Naïve Bayes; TF-IDF.

1. LATAR BELAKANG

Perkembangan teknologi informasi di Indonesia telah mendorong peningkatan pesat pada sektor *e-commerce* dalam beberapa tahun terakhir. Ecommerce didefinisikan sebagai aktivitas jual beli barang maupun jasa melalui jaringan internet dengan dukungan sistem pembayaran dan distribusi yang terintegrasi (Rahmawati, Tamyiz, & Nugroho, 2023). Pergeseran pola konsumsi masyarakat dari konvensional ke arah digital menjadikan *e-commerce* sebagai salah satu sektor strategis dalam era ekonomi digital.

Berdasarkan laporan GoodStats (2023), Lazada menempati posisi ketiga sebagai platform *e-commerce* dengan jumlah pengunjung terbanyak di Indonesia, setelah Shopee dan

Tokopedia. Lazada mencatat 83,2 juta kunjungan pada kuartal I dan 74,5 juta kunjungan pada kuartal II tahun 2023. Tingginya jumlah pengguna tersebut mencerminkan luasnya basis konsumen yang dimiliki Lazada. Namun demikian, mayoritas ulasan pengguna masih terbatas pada rating bintang sehingga informasi yang lebih kaya dari teks ulasan belum dimanfaatkan secara optimal (Aurelia, Sa, Fauzi, & Semarang, 2024).



Gambar 1. Statistik E-commerce Paling Banyak Dikunjungi di Indonesia
Flowchart Kerangka Pemikiran Penelitian 1(Q1–Q2 2023).

Sumber: GoodStats, diolah dari SimilarWeb (2023)

Analisis sentimen hadir sebagai solusi untuk mengolah teks ulasan pengguna agar dapat dikategorikan ke dalam sentimen positif atau negatif. Dengan demikian, perusahaan dapat mengevaluasi kualitas layanan, mengidentifikasi kelemahan, serta menyusun strategi perbaikan berbasis data (Alfaris & Kusnawi, 2023)(Pratmanto, Widayanto, Meisella Kristania, & Wijianto, 2024) Berbagai algoritma *machine learning* telah digunakan untuk analisis sentimen, salah satunya *Naïve Bayes (NB)* yang dikenal sederhana, cepat, dan efektif dalam klasifikasi teks, serta *K-Nearest Neighbor (KNN)* yang bekerja berdasarkan kedekatan antar data sehingga mampu menangkap pola lokal (Aurelia et al., 2024)(Alfaris & Kusnawi, 2023)

Sejumlah penelitian terdahulu telah mengkaji analisis sentimen pada *ecommerce* di Indonesia. Aurelia et al. (2024) menerapkan *Naïve Bayes* pada ulasan Tokopedia dengan hasil akurasi 83,9%. Alfaris dan Kusnawi (2023) membandingkan *Naïve Bayes* dan KNN pada ulasan Shopee dengan hasil performa yang relatif seimbang. Pratmanto et al. (2024) menemukan bahwa KNN unggul pada analisis ulasan aplikasi Vidio, sedangkan (Calvin Wendy & Ade Maulana, 2024) menunjukkan keunggulan *Naïve Bayes* pada aplikasi Ajaib Kripto. Dari studi tersebut terlihat bahwa performa algoritma sangat bergantung pada

karakteristik dataset yang digunakan. Namun demikian, penelitian yang secara khusus membandingkan kinerja *Naïve Bayes* dan KNN pada ulasan aplikasi Lazada berbahasa Indonesia masih sangat terbatas.

Berdasarkan berbagai penelitian tersebut, terlihat bahwa performa algoritma sangat bergantung pada karakteristik dataset yang digunakan (Aggarwal, 2018). Namun hingga kini, belum ada penelitian yang secara khusus membandingkan performa Naïve Bayes dan K-Nearest Neighbor pada ulasan pengguna aplikasi Lazada berbahasa Indonesia (Manning, Raghavan, & Schütze, 2008). Padahal, perbedaan konteks platform, gaya bahasa pengguna, dan karakteristik layanan dapat memengaruhi hasil analisis sentimen (Liu, 2020).

Oleh karena itu, penelitian ini berupaya mengisi kesenjangan tersebut dengan menganalisis dan membandingkan kinerja kedua algoritma untuk menentukan metode yang paling akurat dalam mengklasifikasikan sentimen ulasan aplikasi Lazada. Berdasarkan hal tersebut, penelitian ini dilakukan untuk mengisi kesenjangan dengan menganalisis dan membandingkan kinerja algoritma Naïve Bayes dan KNearest Neighbor pada ulasan aplikasi Lazada yang dikumpulkan dari Google Play Store pada Agustus–September 2023 (Kurniawan & Santoso, 2021).

2. KAJIAN TEORITIS

E-commerce

E-commerce merupakan aktivitas perdagangan berbasis internet yang mengintegrasikan proses transaksi, pembayaran, distribusi, dan layanan purna jual dalam satu ekosistem digital (Rahmawati, Tamyiz, & Nugroho, 2023). Perkembangan *e-commerce* di Indonesia didorong oleh meningkatnya penetrasi internet dan penggunaan perangkat mobile, dengan platform seperti Shopee, Tokopedia, dan Lazada sebagai pemain utama.

Analisis Sentimen

Analisis sentimen merupakan bagian dari *Natural Language Processing* (NLP) yang bertujuan mengidentifikasi dan mengklasifikasikan opini atau emosi dalam teks ke dalam kategori sentimen tertentu, seperti positif, negatif, atau netral (Liu, 2020). Dalam konteks *e-commerce*, analisis sentimen digunakan untuk memahami persepsi dan tingkat kepuasan pelanggan terhadap produk maupun layanan, serta mendukung evaluasi kualitas layanan dan pengambilan keputusan berbasis data (Kadir & Fairuzabadi, 2025).

Natural Language Processing (NLP)

Natural Language Processing (NLP) merupakan cabang kecerdasan buatan yang berfokus pada pemrosesan dan pemahaman bahasa alami manusia oleh komputer (Jurafsky &

Martin, 2021). Dalam analisis sentimen, NLP berperan sebagai fondasi utama untuk mengolah data teks tidak terstruktur, seperti ulasan pengguna aplikasi. Tahapan penting dalam NLP meliputi *preprocessing* teks, seperti pembersihan data, tokenisasi, penghapusan *stopwords*, dan *stemming*, yang bertujuan meningkatkan kualitas representasi teks sebelum digunakan dalam pemodelan klasifikasi sentimen.

Algoritma Naïve Bayes

Naïve Bayes merupakan algoritma klasifikasi berbasis probabilistik yang didasarkan pada Teorema Bayes dengan asumsi independensi antar fitur (Manning et al., 2020). Meskipun asumsi tersebut bersifat sederhana, *Naïve Bayes* terbukti efektif dalam klasifikasi teks berdimensi tinggi karena efisien secara komputasi dan mudah diimplementasikan.

Algoritma K-Nearest Neighbor (KNN)

K-Nearest Neighbor (KNN) merupakan algoritma klasifikasi non-parametrik yang menentukan kelas suatu data berdasarkan kedekatannya dengan k data terdekat dalam ruang fitur (Limbong et al., 2022). Algoritma ini menggunakan ukuran jarak, seperti *Euclidean Distance* atau *Cosine Similarity*, untuk menilai kemiripan antar data teks. KNN unggul dalam menangkap pola lokal dan variasi semantik antar dokumen, namun kinerjanya sensitif terhadap pemilihan nilai k dan ukuran dataset, sehingga relevan untuk dibandingkan dengan algoritma probabilistik seperti *Naïve Bayes* dalam analisis sentimen.

Perkembangan dan Peran Analisis Sentimen di Industri E-commerce

Analisis sentimen merupakan pendekatan komputasional yang berperan penting dalam mendukung pengambilan keputusan berbasis data di industri *e-commerce* dengan mengidentifikasi dan mengukur opini emosional pelanggan dari teks ulasan daring (Liu, 2020). Ulasan pengguna mencerminkan pengalaman nyata terkait kualitas produk dan layanan, sehingga dapat dimanfaatkan untuk memahami persepsi konsumen secara real-time (Rahmawati et al., 2023). Melalui analisis sentimen, perusahaan mampu memantau reputasi merek, mengidentifikasi potensi permasalahan layanan, serta berfungsi sebagai sistem deteksi dini (*early warning system*) dalam menjaga kepuasan dan loyalitas pelanggan di lingkungan digital yang kompetitif.

Tantangan Analisis Sentimen Berbahasa Indonesia

Analisis sentimen berbahasa Indonesia menghadapi tantangan utama berupa karakteristik teks ulasan yang bersifat informal, seperti penggunaan singkatan, emotikon, serta campuran bahasa Indonesia dan Inggris, sebagaimana ditemukan pada ulasan pengguna aplikasi Lazada di Google Play Store. Ulasan ini merefleksikan persepsi pengguna terhadap kualitas layanan dan pengalaman berbelanja, sehingga banyak digunakan dalam penelitian analisis sentimen

karena mampu menggambarkan tingkat kepuasan pelanggan secara dinamis (Abdillah et al., 2024; Alfaris & Kusnawi, 2023). Untuk mengatasi tantangan tersebut, diperlukan tahap preprocessing teks yang komprehensif, meliputi pembersihan karakter tidak relevan, normalisasi kata, tokenisasi, penghapusan stopwords, dan stemming, agar data teks siap digunakan dalam pemodelan klasifikasi sentimen (Aurelia et al., 2024).

Kombinasi dan Pengembangan Metode Analisis Sentimen

Perkembangan *machine learning* mendorong pengembangan metode analisis sentimen melalui penggabungan algoritma klasik seperti *Naïve Bayes* (NB) dan *K-Nearest Neighbor* (KNN) dengan teknik representasi teks yang lebih kompleks. Kombinasi pembobotan *Term Frequency–Inverse Document Frequency* (TF-IDF) dan *word embedding* seperti *Word2Vec* atau *GloVe* terbukti meningkatkan pemahaman konteks semantik kata sehingga menghasilkan klasifikasi sentimen yang lebih akurat (Aurelia et al., 2024). Namun, penerapan pendekatan tersebut pada bahasa Indonesia masih relatif terbatas akibat keterbatasan korpus dan kompleksitas linguistik, sehingga algoritma klasik berbasis *TF-IDF* tetap relevan untuk dianalisis dan dibandingkan kinerjanya.

Penerapan Machine Learning dalam Analisis Opini Publik

Perkembangan teknologi *machine learning* memungkinkan sistem komputer mempelajari pola dari data dan menghasilkan prediksi tanpa pemrograman eksplisit, sehingga banyak dimanfaatkan dalam analisis opini publik di berbagai bidang sosial dan ekonomi (Han, Kamber, & Pei, 2011). Dalam konteks *e-commerce*, pendekatan ini efektif untuk menganalisis sentimen pelanggan yang diekspresikan melalui ulasan daring dan platform digital, dengan mengubah data teks tidak terstruktur menjadi informasi bermakna bagi pengambilan keputusan berbasis data (*data-driven decision making*) (Liu, 2020). Algoritma seperti *Naïve Bayes* dan *K-Nearest Neighbor* (KNN) banyak digunakan karena sederhana, efisien secara komputasi, dan mampu mengklasifikasikan opini pengguna dalam skala besar dengan akurasi yang baik, sehingga membantu perusahaan mengidentifikasi aspek layanan yang paling memengaruhi kepuasan pelanggan dan mendukung perumusan strategi peningkatan kualitas layanan.

Evaluasi Kinerja Model

Evaluasi kinerja model merupakan tahap penting dalam penelitian machine learning untuk menilai kemampuan algoritma dalam menghasilkan prediksi yang akurat dan reliabel. Evaluasi dilakukan menggunakan confusion matrix yang menghasilkan metrik akurasi, presisi, recall, dan F1-score (Sokolova & Lapalme, 2009). Penggunaan lebih dari satu metrik diperlukan karena data sentimen sering kali tidak seimbang antara kelas positif dan negatif. Akurasi memberikan gambaran umum performa model, sedangkan presisi dan recall menilai

ketepatan serta kelengkapan prediksi kelas tertentu. F1-score digunakan untuk memberikan ukuran yang lebih seimbang terhadap kinerja model secara keseluruhan.

Penelitian Relevan

Sejumlah penelitian terdahulu telah mengkaji analisis sentimen pada platform *e-commerce* menggunakan algoritma *Naïve Bayes* (NB) dan *K-Nearest Neighbor* (KNN) dengan memanfaatkan data ulasan berbahasa Indonesia. Aurelia et al. (2024) menerapkan NB dengan pembobotan TF-IDF pada ulasan Tokopedia dan memperoleh akurasi tinggi, meskipun masih menghadapi kendala dalam menangani teks informal. Selain itu, Alfaris dan Kusnawi (2023) serta Limbong et al. (2022) membandingkan kinerja NB dan KNN pada ulasan Shopee dan menunjukkan bahwa NB lebih stabil dan efisien secara komputasi, sedangkan KNN sensitif terhadap pemilihan nilai parameter k dan kurang optimal pada dataset berukuran besar.

Penelitian lain juga menegaskan keunggulan relatif kedua algoritma tersebut pada konteks yang berbeda. Calvin Wendy dan Ade Maulana (2024) menemukan bahwa NB lebih unggul dalam kecepatan dan efisiensi pemrosesan dibandingkan KNN pada ulasan aplikasi Ajaib Kripto, sementara Rahmawati, Tamyiz, dan Nugroho (2023) menunjukkan efektivitas NB dalam menganalisis opini publik e-commerce melalui data Twitter berbahasa Indonesia yang bersifat pendek dan informal. Berdasarkan temuan tersebut, dapat disimpulkan bahwa NB dan KNN efektif untuk klasifikasi sentimen teks berbahasa Indonesia, namun sebagian besar penelitian sebelumnya belum berfokus pada ulasan pengguna Lazada dan belum mengeksplorasi sensitivitas parameter k pada KNN secara mendalam. Oleh karena itu, penelitian ini mengisi celah penelitian dengan menerapkan kedua algoritma tersebut pada ulasan aplikasi Lazada serta menganalisis variasi nilai k untuk menilai pengaruhnya terhadap performa klasifikasi sentimen.

3. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode komparatif karena analisis difokuskan pada pengolahan data numerik yang bersumber dari teks ulasan pengguna aplikasi Lazada. Data teks terlebih dahulu melalui tahapan *text preprocessing* dan *feature extraction* menggunakan teknik *Term Frequency–Inverse Document Frequency* (TF-IDF) untuk mengubah data kualitatif menjadi representasi numerik yang siap dianalisis secara statistik. Pendekatan kuantitatif dipilih karena memungkinkan peneliti memperoleh hasil yang objektif, terukur, dan dapat diuji secara empiris, sebagaimana dinyatakan oleh Sugiyono (2021) bahwa pendekatan ini tepat digunakan untuk penelitian yang berorientasi pada generalisasi berbasis data numerik. Selanjutnya, metode komparatif diterapkan untuk membandingkan

kinerja dua algoritma *machine learning*, yaitu *Naïve Bayes* (NB) dan *K-Nearest Neighbor* (KNN), dalam mengklasifikasikan sentimen ulasan pengguna berdasarkan metrik akurasi, presisi, recall, dan F1-score, sekaligus menilai efisiensi komputasi, stabilitas performa, serta sensitivitas parameter, khususnya variasi nilai *k* pada algoritma KNN.

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh melalui teknik *web scraping* dari laman Google Play Store pada aplikasi Lazada Indonesia selama periode Januari hingga Maret 2024. Google Play Store dipilih sebagai sumber data karena menyediakan ulasan langsung dari pengguna aktif yang merepresentasikan pengalaman nyata pelanggan, baik dalam bentuk teks ulasan maupun rating bintang sebagai indikator awal arah sentimen. Proses pengumpulan data dilakukan menggunakan bahasa pemrograman Python dengan memanfaatkan pustaka *BeautifulSoup* dan *Selenium* pada lingkungan Google Colab, sehingga memungkinkan ekstraksi data dilakukan secara otomatis dan efisien. Sebanyak 10.000 ulasan berbahasa Indonesia berhasil dihimpun dan selanjutnya melalui tahap verifikasi serta *data cleaning* untuk menghapus duplikasi, teks kosong, dan ulasan non-Indonesia guna menjaga kualitas data. Data yang telah dibersihkan disimpan dalam format CSV dan dianalisis lebih lanjut menggunakan Python versi 3.10 dengan dukungan pustaka *pandas*, *NLTK*, *Sastrawi*, dan *scikit-learn* untuk prapemrosesan teks, pembobotan fitur, pelatihan model NB dan KNN, serta evaluasi kinerja model, sehingga seluruh tahapan penelitian dapat dilaksanakan secara sistematis dan dapat direplikasi.

4. HASIL DAN PEMBAHASAN

Implementasi

Proses implementasi penelitian ini dilakukan menggunakan bahasa pemrograman Python yang dikenal luas dalam bidang *machine learning* dan *natural language processing (NLP)* karena memiliki pustaka (*library*) pendukung yang lengkap serta mudah digunakan. Beberapa pustaka utama yang dimanfaatkan meliputi *scikit-learn*, *pandas*, *numpy*, *matplotlib*, dan *Sastrawi*. Pustaka *scikit-learn* digunakan untuk keperluan klasifikasi, pembagian data, evaluasi model, serta penerapan metode *10-Fold Cross Validation*. Pustaka *pandas* berfungsi untuk membaca, mengelola, dan memanipulasi dataset dalam format CSV. Pustaka *numpy* digunakan untuk operasi *numerik* dan perhitungan matriks, sementara *matplotlib* dimanfaatkan untuk memvisualisasikan hasil evaluasi, seperti *Confusion Matrix*. Adapun *Sastrawi* digunakan untuk proses *stopword removal* dan *stemming* terhadap teks berbahasa Indonesia dengan spesifikasi sebagai berikut:

Tabel 1. Proses *Stopword Removal* dan *Stemming* Terhadap Teks.

Komponen	Spesifikasi
Sistem Operasi	Windows 11 64-bit
Prosesor	Intel Core i5
RAM	8 GB
Bahasa	Python 3.10
Pemrograman	
Library Utama	scikit-learn, pandas, numpy, matplotlib, Sastrawi

Tahapan implementasi model meliputi:

a. Pra-pemrosesan Teks (*Text Preprocessing*)

Tahap ini meliputi proses *cleansing*, *case folding*, *tokenisasi*, *stopword removal*, dan *stemming* untuk memastikan teks bersih dan seragam sebelum dilakukan analisis.

b. Representasi Data Teks (*Feature Extraction*)

Setelah tahap pembersihan, teks diubah menjadi bentuk numerik menggunakan dua metode, yaitu Term *Frequency–Inverse Document Frequency* (TF-IDF) dan *CountVectorizer*, agar dapat diproses oleh algoritma klasifikasi.

c. Pelatihan Model (*Model Training*)

Dua algoritma pembelajaran mesin diterapkan, yaitu *Naïve Bayes (NB)* dan *K-Nearest Neighbor (KNN)*. Kedua algoritma ini dipilih karena memiliki performa yang baik pada data teks berbahasa Indonesia dan sering digunakan dalam penelitian analisis sentimen.

d. Pengujian Model (*Model Testing*)

Model diuji menggunakan metode *10-Fold Cross Validation* untuk memastikan hasil evaluasi lebih stabil dan representatif terhadap seluruh dataset.

e. Evaluasi Kinerja Model (*Model Evaluation*)

Kinerja model diukur menggunakan empat metrik utama, yaitu akurasi, presisi, *recall*, dan *F1-score*, yang digunakan untuk menilai seberapa baik model mengenali dan mengklasifikasikan sentimen positif maupun negatif.

Hasil Eksperimen

Kinerja Naïve Bayes

Naïve Bayes dilatih menggunakan *CountVectorizer* untuk representasi teks. Model menghasilkan performa yang sangat baik dengan akurasi 90,05%, presisi 89,93%, recall 90,05%, dan F1-score 89,96%. Detail per kelas menunjukkan bahwa pada kategori positif, nilai presisi mencapai 0,92 dengan *recall* 0,94, sementara pada kategori negatif, nilai presisi 0,86 dan recall 0,80. Hal ini menandakan bahwa *Naïve Bayes* mampu mengenali ulasan positif dengan sangat baik, namun sedikit menurun pada prediksi ulasan negatif. Hasil ini sejalan

dengan penelitian sebelumnya, di mana *Naïve Bayes* terbukti efektif untuk klasifikasi teks karena kesederhanaan dan kecepatan komputasinya (Alamsyah & Saviera, 2021; Susanti & Ilahi, 2024).

Pengaruh Variasi Nilai k terhadap Kinerja KNN

Untuk mengoptimalkan performa algoritma *K-Nearest Neighbor*, penelitian ini juga melakukan eksperimen variasi nilai parameter k dengan mencoba beberapa nilai yaitu k = 3, 5, 7, dan 9. Pengujian dilakukan menggunakan metrik jarak *Cosine Similarity*, karena metode ini dinilai lebih sesuai untuk data teks dibandingkan *Euclidean Distance*.

Hasil pengujian menunjukkan bahwa nilai k = 5 memberikan performa terbaik dengan keseimbangan antara akurasi, presisi, recall, dan *F1-score*. Nilai k yang terlalu kecil (misalnya k = 3) membuat model menjadi terlalu sensitif terhadap *noise*, sehingga cenderung mengalami *overfitting*. Sebaliknya, nilai k yang terlalu besar (misalnya k = 9) membuat model kehilangan kemampuan membedakan pola lokal antar data dan berpotensi menghasilkan prediksi yang kurang spesifik. Tabel berikut merangkum hasil pengujian variasi nilai k:

Tabel 2. Hasil Pengujian Variasi Nilai K.

Nilai K	Akurasi	Presisi	Recall	F1-score
3	71,00%	73,00%	71,00%	63,00%
5	73,00%	75,00%	73,00%	65,00%
7	71,80%	73,20%	71,80%	64,00%
9	70,50%	71,90%	70,50%	63,10%

Dari tabel di atas, terlihat bahwa nilai K = 5 menghasilkan performa terbaik dengan:

- a. Akurasi = 73,00%
- b. Presisi = 75,00%
- c. Recall = 73,00%
- d. F1-score = 65,00%

Nilai ini menunjukkan bahwa pada K = 5, model memiliki keseimbangan terbaik antara kemampuan mengenali ulasan positif dan negatif. Nilai K yang terlalu kecil (misalnya K=3) membuat model terlalu sensitif terhadap data latih (*overfitting*), sedangkan nilai K yang terlalu besar (misalnya K=9) membuat model kehilangan kemampuan membedakan pola secara spesifik (*underfitting*). Dengan demikian, dapat disimpulkan bahwa K=5 merupakan nilai optimal untuk algoritma KNN dalam penelitian ini, karena memberikan kombinasi performa terbaik di antara seluruh pengujian yang dilakukan.

Perbandingan Kinerja Algoritma

Perbandingan Bentuk Tabel

Tabel 2 menampilkan hasil pengujian kinerja kedua algoritma, yaitu *Naïve Bayes* dan *K-Nearest Neighbor (KNN)*, berdasarkan empat metrik evaluasi: akurasi, presisi, *recall*, dan *F1-score*.

Tabel 3. Perbandingan Kinerja Algoritma *Naïve Bayes* dan *K-Nearest Neighbor*.

Algoritma	Akurasi	Presisi	Recal	F1-score
<i>Naïve Bayes</i>	0,89%	0,89%	0,89%	0,89%
KNN	0,73%	0,75%	0,73%	0,65%

Sumber: Hasil Olahan Data Penelitian (2024)

Proses evaluasi dilakukan menggunakan metode *10-Fold Cross Validation* untuk memastikan hasil yang diperoleh bersifat stabil dan tidak bergantung pada satu pembagian data tertentu. Empat metrik evaluasi utama digunakan, yaitu akurasi, presisi, *recall*, dan *F1-score*. Hasil pengujian kedua algoritma, yaitu *Naïve Bayes* dan *K-Nearest Neighbor (KNN)*, dapat dilihat pada tabel di atas.

1) Akurasi

Nilai akurasi model *Naïve Bayes* sebesar 0,8956 (89,56%) menunjukkan bahwa model mampu memprediksi sentimen dengan benar pada hampir 9 dari 10 data ulasan. Artinya, sebagian besar prediksi yang dilakukan model sudah sesuai dengan label sentimen sebenarnya .Sementara itu, algoritma KNN memperoleh akurasi 0,7322 (73,22%), yang berarti hanya sekitar 7 dari 10 prediksi yang tepat. Hal ini menunjukkan bahwa *Naïve Bayes* memiliki tingkat ketepatan klasifikasi yang lebih tinggi dibanding KNN dalam mengenali sentimen positif dan negatif pada data ulasan Lazada.

2) Presisi

Presisi pada *Naïve Bayes* mencapai 0,8953 (89,53%), yang mengindikasikan bahwa dari seluruh prediksi positif yang dihasilkan oleh model, sekitar 89,53% benar-benar merupakan ulasan positif. Dengan kata lain, tingkat kesalahan dalam mengklasifikasikan ulasan negatif sebagai positif relative kecil. Sebaliknya, KNN memiliki presisi 0,7521 (75,21%), yang masih cukup baik, namun menunjukkan bahwa model ini cenderung sedikit lebih sering salah dalam memprediksi *sentimen* positif. Hal ini bisa terjadi karena KNN sensitif terhadap distribusi data dan jarak antar titik, terutama jika data tidak terdistribusi seimbang.

3) Recall

Recall pada *Naïve Bayes* sebesar 0,8956 (89,56%), menunjukkan kemampuan model dalam mengenali seluruh ulasan positif dari data yang sebenarnya positif. Nilai *recall* yang tinggi menandakan bahwa *Naïve Bayes* memiliki kemampuan mendeteksi sentimen positif secara menyeluruh. Pada sisi lain, KNN memperoleh *recall* 0,7322 (73,22%), yang berarti masih ada sebagian ulasan positif yang tidak berhasil dideteksi dengan benar dan diklasifikasikan sebagai negatif. Hal ini menunjukkan bahwa KNN kurang optimal dalam menjangkau seluruh variasi pola teks pada dataset yang besar.

4) F1-score

F1-score pada *Naïve Bayes* mencapai 0,8953 (89,53%), yang merupakan hasil keseimbangan antara presisi dan *recall*. Nilai yang tinggi ini membuktikan bahwa model *Naïve Bayes* tidak hanya akurat, tetapi juga konsisten dalam menjaga keseimbangan antara ketepatan dan kelengkapan prediksi. Sebaliknya, KNN memperoleh nilai 0,6564 (65,64%), menandakan bahwa meskipun presisinya masih cukup baik, *recall* yang relatif rendah menyebabkan keseimbangan performa model menurun. Hal ini menegaskan bahwa *Naïve Bayes* lebih stabil dan efisien dalam menangani data teks berbahasa Indonesia yang bersifat bervariasi dan panjang.

Berdasarkan hasil evaluasi, dapat disimpulkan bahwa algoritma *Naïve Bayes* memberikan performa terbaik dengan rata-rata metrik di atas 89%, jauh lebih unggul dibandingkan KNN yang hanya mencapai sekitar 73% akurasi. Performa *Naïve Bayes* yang lebih baik dapat dijelaskan karena algoritma ini mampu mengasumsikan independensi antar kata (fitur), sehingga lebih efisien dalam memproses teks dalam jumlah besar. Sedangkan KNN cenderung tidak stabil terhadap ukuran data dan rentan terhadap noise pada data teks. Dengan demikian, *Naïve Bayes* merupakan algoritma yang paling optimal untuk analisis sentimen ulasan pengguna aplikasi Lazada pada penelitian ini.

Perbandingan Bentuk Diagram

Gambar 1 berikut menampilkan perbandingan kinerja kedua algoritma dalam bentuk diagram batang berdasarkan keempat metrik evaluasi yang sama.



Gambar 2. Perbandingan Kinerja Algoritma *Naïve Bayes* dan KNN pada Analisis Sentimen Ulasan Lazada.

Sumber: Hasil Olahan Data Penelitian (2024)

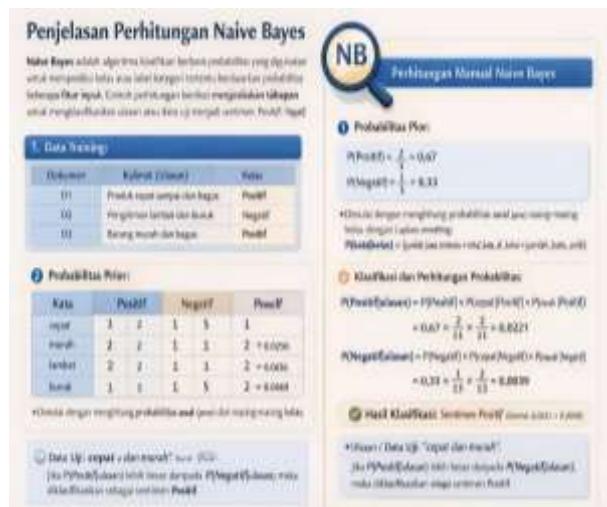
Dari visualisasi pada Gambar 1 dapat dilihat bahwa *Naïve Bayes* unggul secara konsisten pada semua metrik dibandingkan KNN. Nilai akurasi *Naïve Bayes* sebesar 90,05%, jauh lebih tinggi dibandingkan 71,90% pada KNN, menunjukkan adanya perbedaan performa yang signifikan.

Selain itu, *F1-score* *Naïve Bayes* sebesar 89,96% menunjukkan keseimbangan yang baik antara presisi dan recall, sementara KNN hanya mencapai 62,53%, yang menegaskan bahwa KNN lebih sensitif terhadap distribusi data yang tidak seimbang.

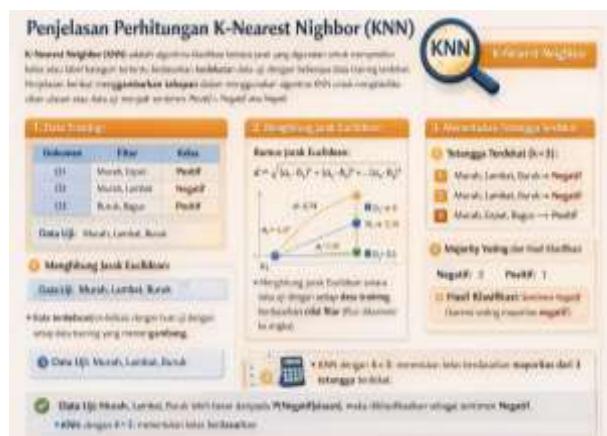
Temuan ini memperkuat bahwa *Naïve Bayes* lebih stabil, efisien, dan akurat dalam menangani data teks berbahasa Indonesia yang kompleks seperti ulasan aplikasi Lazada. Secara umum, hasil ini mengonfirmasi bahwa *Naïve Bayes* merupakan algoritma yang paling optimal untuk analisis sentimen ulasan Lazada, baik dari segi akurasi maupun efisiensi komputasi.

Perhitungan Manual Kinerja Algoritma

Untuk memastikan pemahaman terhadap proses klasifikasi yang dilakukan sistem, dilakukan perhitungan manual sederhana terhadap masing-masing algoritma, yaitu *Naïve Bayes* dan *K-Nearest Neighbor (KNN)*. Contoh berikut menggunakan data ulasan singkat untuk menggambarkan langkah-langkah perhitungan probabilitas dan jarak antar data.



Gambar 3. Perhitungan Manual *Naïve Bayes*.



Gambar 4. Perhitungan Manual *K-Nearest Neighbor (KNN)*.

Berdasarkan hasil perhitungan manual, kedua algoritma menghasilkan klasifikasi yang sama, yaitu sentimen positif untuk data uji “pengiriman cepat”. Hal ini sejalan dengan hasil pengujian sistem yang menunjukkan bahwa *Naïve Bayes* dan KNN memiliki kemampuan serupa dalam mengenali pola sentimen positif, meskipun akurasi keseluruhan *Naïve Bayes* lebih tinggi pada dataset penelitian.

Pembahasan

Berdasarkan hasil eksperimen yang dilakukan terhadap dataset ulasan aplikasi Lazada, algoritma *Naïve Bayes* menunjukkan performa yang lebih unggul dibandingkan dengan *K-Nearest Neighbor (KNN)* pada seluruh metrik evaluasi, yaitu akurasi, presisi, *recall*, dan *F1-score*. *Naïve Bayes* memperoleh akurasi sebesar 90,05%, sedangkan KNN hanya mencapai 71,90%. Perbedaan performa yang cukup signifikan ini menunjukkan bahwa *Naïve Bayes* lebih sesuai digunakan untuk analisis sentimen teks ulasan berbahasa Indonesia yang bersifat kompleks dan berdimensi tinggi.

Selain itu, *Naïve Bayes* tidak memerlukan proses tuning parameter, seperti halnya pada KNN yang sangat bergantung pada nilai k dan pemilihan metrik jarak. Tanpa adanya parameter yang perlu dioptimalkan, *Naïve Bayes* dapat langsung menghasilkan hasil klasifikasi yang konsisten dan stabil. Di sisi lain, KNN memerlukan pencarian nilai k terbaik, dan performanya sangat sensitif terhadap distribusi data serta keberadaan data yang tidak seimbang (imbalanced dataset). Dalam penelitian ini, meskipun telah dilakukan *oversampling* pada kelas minoritas, KNN masih cenderung bias terhadap kelas mayoritas (ulasan positif), yang terlihat dari recall kelas negatif yang hanya sebesar 0,08, menunjukkan lemahnya kemampuan KNN dalam mengenali ulasan negatif.

Faktor lain yang berkontribusi pada keunggulan *Naïve Bayes* adalah efisiensi komputasi. Algoritma ini melakukan perhitungan probabilitas secara langsung tanpa harus menghitung jarak antar setiap pasangan data seperti yang dilakukan KNN. Dengan demikian, waktu eksekusi *Naïve Bayes* jauh lebih singkat, menjadikannya lebih efisien ketika diterapkan pada jumlah data yang besar, seperti 10.000 ulasan pengguna Lazada dalam penelitian ini.

Secara keseluruhan, keunggulan *Naïve Bayes* dibandingkan KNN dapat dijelaskan melalui empat aspek utama:

- a. Efisiensi pada data berdimensi tinggi: *Naïve Bayes* mampu mengolah ribuan fitur kata dengan cepat tanpa penurunan performa signifikan.
- b. Stabilitas terhadap distribusi data tidak seimbang: *Naïve Bayes* tetap memberikan hasil klasifikasi yang baik meskipun proporsi kelas tidak seimbang.
- c. Tidak membutuhkan tuning parameter: Berbeda dengan KNN, performa *Naïve Bayes* tidak bergantung pada parameter seperti nilai k .
- d. Kecepatan proses pelatihan dan prediksi: *Naïve Bayes* memiliki waktu komputasi yang jauh lebih singkat dibandingkan KNN karena tidak menghitung jarak antar seluruh data.

Temuan ini konsisten dengan penelitian sebelumnya oleh Calvin Wendy dan Ade Maulana (2024), yang menyimpulkan bahwa *Naïve Bayes* lebih cepat dan stabil dibandingkan KNN pada analisis sentimen aplikasi Ajaib Cripto. Dengan demikian, hasil penelitian ini memperkuat bukti empiris bahwa *Naïve Bayes* merupakan algoritma yang lebih efisien dan akurat untuk analisis sentimen ulasan pelanggan dalam konteks *e-commerce* di Indonesia.

Berdasarkan hasil pengujian yang telah dilakukan, algoritma *Naïve Bayes* menunjukkan performa yang lebih unggul dibandingkan *K-Nearest Neighbor (KNN)* dalam mengklasifikasikan sentimen ulasan pengguna terhadap aplikasi Lazada. Hal ini menunjukkan bahwa pendekatan probabilistik yang digunakan oleh *Naïve Bayes* lebih efektif dalam

menangani teks pendek dengan variasi kata yang tinggi, seperti ulasan pengguna di platform e-commerce.

Hasil penelitian ini memberikan implikasi praktis yang dapat dimanfaatkan oleh pihak Lazada untuk meningkatkan kualitas layanan dan pengalaman pelanggan. Penerapan model analisis sentimen berbasis *Naïve Bayes* dapat digunakan sebagai alat bantu otomatis untuk memantau sentimen pelanggan secara berkelanjutan. Melalui sistem ini, manajemen dapat dengan cepat mengidentifikasi aspek-aspek layanan yang sering mendapat keluhan, seperti keterlambatan pengiriman, kualitas produk, maupun respons dari layanan pelanggan.

Selain itu, hasil analisis sentimen dapat menjadi dasar bagi perusahaan dalam merumuskan strategi peningkatan layanan. Misalnya, ketika sentimen negatif mendominasi pada kategori pengiriman, Lazada dapat segera melakukan evaluasi terhadap sistem logistik atau memperkuat kerja sama dengan mitra ekspedisi agar proses distribusi lebih cepat dan efisien. Dengan memahami pola sentimen pengguna secara lebih mendalam, Lazada juga dapat meningkatkan kepercayaan pelanggan melalui tindakan responsif terhadap keluhan yang muncul. Hal ini pada akhirnya akan berkontribusi terhadap peningkatan kepuasan dan loyalitas pengguna, serta memperkuat citra positif perusahaan di tengah persaingan pasar *e-commerce* yang semakin ketat.

Lebih jauh lagi, penerapan analisis sentimen ini dapat mendukung proses pengambilan keputusan berbasis data (*data-driven decision making*) di lingkungan manajerial Lazada. Hasil klasifikasi sentimen dapat dijadikan bahan evaluasi terhadap efektivitas kampanye promosi, peluncuran fitur baru, maupun kebijakan pelayanan yang diterapkan. Dengan memanfaatkan hasil analisis ini secara konsisten, Lazada dapat membuat keputusan strategis yang lebih tepat sasaran, objektif, dan berbasis pada persepsi nyata pengguna.

Secara keseluruhan, penelitian ini tidak hanya memberikan kontribusi dalam ranah akademik, tetapi juga menawarkan nilai praktis yang tinggi bagi perusahaan *e-commerce* seperti Lazada, terutama dalam memahami opini pelanggan dan meningkatkan kualitas hubungan antara perusahaan dan pengguna.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil perbandingan kinerja algoritma *Naïve Bayes* dan *K-Nearest Neighbor* (KNN) dalam analisis sentimen ulasan aplikasi Lazada, dapat disimpulkan bahwa *Naïve Bayes* menunjukkan performa yang lebih optimal dengan capaian akurasi 90,05%, presisi 89,93%, *recall* 90,05%, dan *F1-score* 89,96%, yang mencerminkan keseimbangan antara ketepatan prediksi dan stabilitas model. Sebaliknya, algoritma KNN menghasilkan kinerja yang lebih

rendah dengan akurasi 71,90% dan *F1-score* 62,53%, yang mengindikasikan kesulitan dalam mengklasifikasikan sentimen negatif akibat sensitivitas terhadap ketidakseimbangan data dan variasi kata. Meskipun KNN berpotensi bekerja lebih baik pada data yang seimbang dengan pemilihan parameter k yang optimal, secara keseluruhan *Naïve Bayes* lebih efektif dan efisien untuk diterapkan dalam analisis sentimen ulasan pengguna aplikasi Lazada.

Beberapa saran untuk penelitian selanjutnya mencakup penggunaan teknik penyeimbangan data, seperti SMOTE atau undersampling, untuk meningkatkan performa algoritma KNN agar tidak bias terhadap kelas mayoritas, serta eksplorasi berbagai metode representasi fitur, seperti Word2Vec, GloVe, atau FastText, sebagai perbandingan dengan TF-IDF dan CountVectorizer yang digunakan dalam penelitian ini. Selain itu, untuk meningkatkan akurasi prediksi, penelitian mendatang dapat mengembangkan model berbasis Deep Learning, seperti LSTM atau Bidirectional LSTM, yang mampu menangkap konteks kata secara lebih mendalam. Hasil penelitian juga berpotensi dikembangkan menjadi platform monitoring sentimen otomatis yang membantu perusahaan menganalisis persepsi pengguna secara real-time. Dengan pengembangan tersebut, penelitian diharapkan dapat memberikan kontribusi signifikan dalam pengembangan sistem analisis sentimen yang lebih efektif, akurat, dan adaptif terhadap berbagai jenis data teks ulasan di masa mendatang.

DAFTAR REFERENSI

- 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), 32–44. <https://doi.org/10.33998/processor.2024.19.1.1596>
- Aggarwal, C. C. (2018). *Machine learning for text*. Springer. <https://doi.org/10.1007/978-3-319-73531-3>
- Alamsyah, A., & Saviera, F. (n.d.). *A comparison of Indonesia's e-commerce sentiment analysis for marketing intelligence effort (case study of Bukalapak, Tokopedia and Elevenia)* (pp. 1–6).
- Alfaris, S., & Kusnawi. (2023). Komparasi metode KNN dan Naive Bayes terhadap analisis sentimen pengguna aplikasi Shopee. *Indonesian Journal of Computer Science*, 12(5), 2766–2776. <https://doi.org/10.33022/ijcs.v12i5.3304>
- Alfarros, M., Naufalino, D., & Al-Husaini, M. (2025). Comparison of support vector machine, random forest, and XGBoost for sentiment analysis on Indodax. *Journal of Computer Networks, Architecture and High Performance Computing*, 7(2), 562–572. <https://doi.org/10.47709/cnahpc.v7i2.5894>
- Aurelia, N., Sa, U., Fauzi, F., & Universitas Negeri Semarang. (2024). Analisis sentimen pada ulasan aplikasi menggunakan klasifikasi Naïve Bayes Tokopedia. *Jurnal Teknologi dan Komputer*, 44–51.

- Calvin, W., & Maulana, A. (2024). Perbandingan analisis sentimen ulasan aplikasi Ajaib Kripto menggunakan metode Naïve Bayes dan K-Nearest Neighbor. *JUMINTAL: Jurnal Manajemen Informatika dan Bisnis Digital*, 3(2), 72–84. <https://doi.org/10.55123/jumintal.v3i2.3965>
- Diastama, A., Firdaus, A., Rahmawan, R. D., Mahendra, Y. R., Cahyono, H. D., & Universitas Sebelas Maret. (2024). Klasifikasi sentimen analisis pada women's e-commerce reviews dengan pendekatan machine learning. *Jurnal Teknik Informatika (JUTIF)*, 5(6), 1549–1559. <https://doi.org/10.52436/1.jutif.2024.5.6.2392>
- Enjelia, L., Cahyana, Y., Rahmat, & Wahiddin, D. (2025). Comparison of K-nearest neighbors and Naive Bayes classifier algorithms in sentiment analysis of 2024 election in Twitter (X). *Journal of Applied Informatics and Computing*, 9(3), 946–954. <https://doi.org/10.30871/jaic.v9i3.9593>
- Kurniawan, D., & Santoso, B. (2021). Analisis sentimen ulasan aplikasi e-commerce menggunakan metode klasifikasi berbasis teks. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 8(4), 789–798. <https://doi.org/10.25126/jtiik.2021841234>
- Liu, B. (2020). *Sentiment analysis: Mining opinions, sentiments, and emotions* (2nd ed.). Cambridge University Press. <https://doi.org/10.1017/9781108639286>
- Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to information retrieval*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511809071>
- Mardiana, N. A., & Windari, W. O. (2024). Analisis sentimen terhadap ulasan e-commerce menggunakan metode machine learning. *G-Tech: Jurnal Teknologi Terapan*, 8(1), 186–195.
- Neighbor, K., Adi, S., Mola, S., Polly, Y. T., & Atok, Y. C. (2025). Analisis sentimen terhadap data komentar publik mengenai isu UU Pilkada 2024 menggunakan metode Naïve Bayes. *Jurnal Ilmu Manajemen dan Akuntansi Terapan (JIMAT)*, 5(3), 337–343. <https://doi.org/10.47065/jimat.v5i3.514>
- Pratmanto, D., Widayanto, A., Kristania, Y. M., & 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.
- Rahmawati, A., Tamyiz, U. M. H., & Nugroho, I. M. (2023). Analisis sentimen tingkat kepopuleran e-commerce pada pengguna Twitter menggunakan algoritma Naive Bayes. *Jurnal Aplikasi dan Teori Ilmu Komputer (JATIKOM)*, 6(1), 21–26. <https://doi.org/10.17509/jatikom.v6i1.48867>