Journal Software, Hardware and Information Technology
p-ISSN: 2808-3385 | e-ISSN: 2776-8961
Volume 6 Issue 1, January 2026

Analisis Sentimen Ulasan Pengguna Aplikasi Jalan Kita 2.0 Menggunakan Algoritma *Naïve Bayes*

Sentiment Analysis of User Reviews of the Jalan Kita 2.0 Application Using the Naïve Bayes Algorithm

Muhammad Ghianza Al Ghifari, Asep Saeppani
Jurusan Teknik Informatika, Universitas Sebelas April, Sumedang, Indonesia
220660121058@student.unsap.ac.id, saeppani@unsap.ac.id

Notes

Submitted: 31-12-2025
Revised: 10-01-2026
Accepted: 17-01-2026
Published: 30-01-2026

How to cite: M. G. Al Ghifari and A. Saeppani, "Analisis Sentimen Ulasan Pengguna Aplikasi Jalan Kita 2.0 Menggunakan Algoritma Naïve Bayes," *Journal Software, Hardware and Information Technology (SHIFT)*, vol. 6, no. 1, pp. 80–91, Jan. 2026, doi: 10.24252/shift.v6i1.271.

DOI: 10.24252/shift.v6i1.271

Copyright 2026© the Author(s)

This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

ABSTRACT

User reviews of the Jalan Kita 2.0 application represent a valuable source of information for evaluating the quality of application-based road infrastructure reporting services. However, these reviews are expressed as unstructured text, which makes manual analysis inefficient and prone to subjectivity. This study aims to analyze the sentiment of user reviews of the Jalan Kita 2.0 application using the Multinomial Naïve Bayes algorithm. The dataset was collected from the Google Play Store and processed through several text preprocessing stages, including case folding, tokenization, stopword removal, and stemming. Feature extraction was performed using the Term Frequency–Inverse Document Frequency (TF-IDF) method. A total of 334 reviews were analyzed and classified into two sentiment categories, namely positive and negative, with the data split into training and testing sets using an 80:20 ratio. The results show that positive sentiment dominates the user reviews, accounting for 81.44% of the dataset, indicating that most users perceive the application as beneficial for reporting road infrastructure issues. Model evaluation achieved an accuracy of 0.93, with precision, recall, and F1-score values of 0.93, 0.93, and 0.92, respectively. These findings demonstrate that the Multinomial Naïve Bayes algorithm performs effectively in classifying sentiment in user reviews of the Jalan Kita 2.0 application and can be utilized as a data-driven basis for evaluating and improving the quality of public service applications.

Keywords: Jalan Kita 2.0, mobile apps, naïve Bayes, sentiment analysis, TF-IDF, user reviews.

ABSTRAK

Ulasan pengguna pada aplikasi Jalan Kita 2.0 merupakan sumber informasi yang bernilai penting dalam mengevaluasi kualitas layanan pelaporan infrastruktur jalan berbasis

aplikasi. Namun, ulasan tersebut umumnya berbentuk teks tidak terstruktur sehingga analisis secara manual menjadi tidak efisien dan berpotensi menimbulkan subjektivitas. Penelitian ini bertujuan untuk menganalisis sentimen ulasan pengguna aplikasi Jalan Kita 2.0 menggunakan algoritma Multinomial Naïve Bayes. Data penelitian diperoleh dari Google Play Store dan diproses melalui tahapan prapemrosesan teks yang meliputi case folding, tokenisasi, stopword removal, dan stemming. Selanjutnya, ekstraksi fitur dilakukan menggunakan metode Term Frequency–Inverse Document Frequency (TF-IDF). Dataset yang digunakan terdiri dari 334 ulasan yang diklasifikasikan ke dalam dua kategori sentimen, yaitu positif dan negatif, serta dibagi menjadi data latih dan data uji dengan rasio 80:20. Hasil penelitian menunjukkan bahwa sentimen positif mendominasi ulasan pengguna, yaitu sebesar 81,44%, yang mengindikasikan bahwa mayoritas pengguna menilai aplikasi ini bermanfaat dalam mendukung pelaporan kondisi infrastruktur jalan. Evaluasi model menghasilkan nilai akurasi sebesar 0,93, dengan nilai precision, recall, dan F1-score masing-masing sebesar 0,93, 0,93, dan 0,92. Temuan ini menunjukkan bahwa algoritma Multinomial Naïve Bayes efektif dalam mengklasifikasikan sentimen ulasan pengguna aplikasi Jalan Kita 2.0 serta dapat dimanfaatkan sebagai dasar evaluasi dan peningkatan kualitas layanan publik berbasis aplikasi.

Kata kunci: analisis sentimen, aplikasi mobile, Jalan Kita 2.0, naïve Bayes, TF-IDF, ulasan pengguna.

1. PENDAHULUAN

Pelayanan publik di bidang infrastruktur jalan memerlukan mekanisme pelaporan yang efektif agar permasalahan di lapangan dapat segera diketahui dan ditindaklanjuti oleh instansi terkait. Aplikasi *Jalan Kita 2.0* hadir sebagai sarana bagi masyarakat untuk menyampaikan laporan dan ulasan terkait kondisi jalan serta proses penanganannya. Ulasan pengguna yang disampaikan melalui aplikasi maupun platform distribusi aplikasi mencerminkan pengalaman langsung masyarakat terhadap layanan yang diberikan, sehingga memiliki nilai strategis sebagai bahan evaluasi layanan publik.

Seiring meningkatnya jumlah ulasan pengguna, data yang dihasilkan berbentuk teks tidak terstruktur dan sulit dianalisis secara manual. Analisis manual tidak hanya membutuhkan waktu dan tenaga yang besar, tetapi juga berpotensi menimbulkan subjektivitas dalam penilaian. Oleh karena itu, diperlukan pendekatan komputasional yang mampu mengolah ulasan pengguna secara sistematis dan objektif. Analisis sentimen menjadi salah satu pendekatan yang banyak digunakan untuk mengidentifikasi kecenderungan opini pengguna terhadap suatu aplikasi atau layanan [1], [2], [3], [4].

Berbagai penelitian sebelumnya telah menerapkan analisis sentimen pada ulasan aplikasi mobile di berbagai domain, seperti transportasi daring, e-commerce, layanan perbankan, pendidikan daring, dan layanan publik. Hasil penelitian menunjukkan bahwa ulasan pengguna dapat

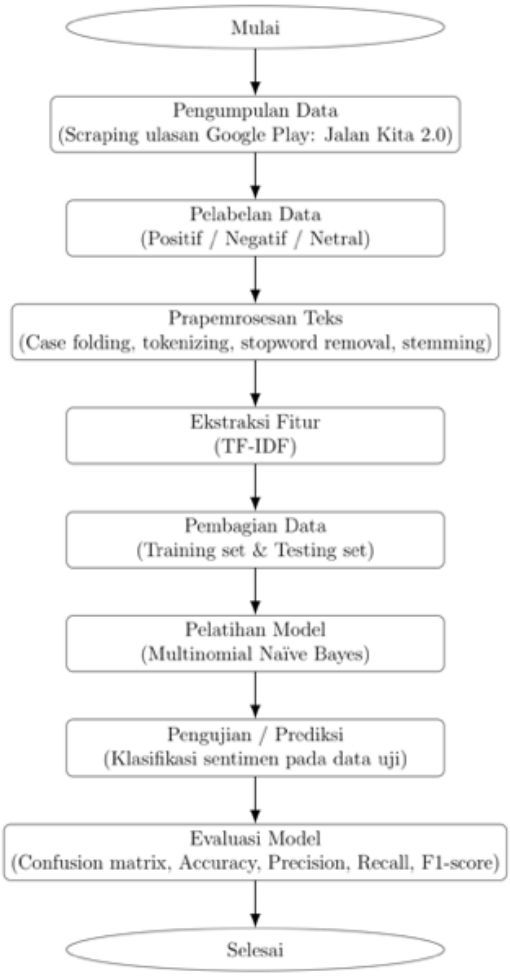
diklasifikasikan secara efektif untuk menggambarkan tingkat kepuasan dan permasalahan layanan yang dirasakan pengguna [5], [6], [7], [8]. Algoritma Naïve Bayes menjadi salah satu metode yang paling banyak digunakan karena kesederhanaan, efisiensi komputasi, dan performanya yang stabil pada klasifikasi teks ulasan aplikasi [9], [10].

Meskipun demikian, sebagian besar penelitian tersebut masih berfokus pada aplikasi komersial atau layanan publik secara umum, seperti aplikasi transportasi, perbankan, dan e-commerce. Penelitian yang secara khusus mengkaji sentimen ulasan pengguna pada aplikasi pelaporan infrastruktur jalan masih relatif terbatas. Selain itu, pemanfaatan hasil analisis sentimen sebagai bahan evaluasi layanan pelaporan jalan berbasis aplikasi belum banyak dibahas secara mendalam dalam penelitian terdahulu.

Berdasarkan celah penelitian tersebut, penelitian ini bertujuan untuk menganalisis sentimen ulasan pengguna aplikasi *Jalan Kita 2.0* menggunakan algoritma Naïve Bayes. Hasil penelitian diharapkan dapat memberikan gambaran kecenderungan opini pengguna serta menjadi dasar evaluasi layanan aplikasi pelaporan infrastruktur jalan berbasis data ulasan pengguna.

2. METODE PENELITIAN

2.1. Jenis dan Pendekatan Penelitian



Gambar 1. Alur Tahapan Penelitian Analisis Sentimen Ulasan Pengguna Aplikasi Jalan Kita 2.0

Alur tahapan penelitian yang digunakan dalam penelitian ini ditunjukkan pada Gambar 1. Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimen pada data teks. Pendekatan kuantitatif dipilih karena tujuan penelitian adalah untuk mengukur dan menguji kinerja algoritma klasifikasi dalam analisis sentimen ulasan pengguna aplikasi Jalan Kita 2.0 berdasarkan data empiris yang bersifat numerik. Proses penelitian terealisasi melalui rangkaian tahapan sistematis mulai dari pengolahan data teks, pelatihan model, hingga evaluasi performa klasifikasi. Pendekatan semacam ini umum digunakan dalam penelitian teks mining untuk memperoleh hasil yang terukur dan dapat diuji secara statistik [11].

2.2. Sumber Data

Sumber data dalam penelitian ini adalah ulasan pengguna dari Google Play Store yang ditulis dalam bahasa Indonesia. Ulasan mencakup teks komentar pengguna serta informasi tambahan seperti tanggal ulasan. Pemilihan data berbahasa Indonesia bertujuan untuk menyelaraskan analisis dengan konteks linguistik pengguna aplikasi di Indonesia, sebagaimana dilakukan dalam penelitian ulasan aplikasi serupa dengan algoritma Naïve Bayes dan ekstraksi fitur TF-IDF [12].

2.3. Implementasi dan Lingkungan Pengembangan

Proses pengolahan data dan analisis sentimen dalam penelitian ini diimplementasikan menggunakan bahasa pemrograman Python dengan lingkungan pengembangan Google Colaboratory (Google Colab). Platform ini dipilih karena menyediakan lingkungan komputasi berbasis cloud yang memadai untuk pemrosesan data teks serta mendukung berbagai pustaka pembelajaran mesin yang umum digunakan. Library utama yang digunakan meliputi pandas dan NumPy untuk pengolahan data, scikit-learn untuk ekstraksi fitur TF-IDF dan klasifikasi Naïve Bayes, serta Sastrawi untuk proses stemming bahasa Indonesia. Penggunaan library tersebut memungkinkan proses prapemrosesan, ekstraksi fitur, pelatihan model, dan evaluasi dilakukan secara terintegrasi dan efisien.

2.4. Pengumpulan Data

Tabel 1. Contoh Data Mentah Ulasan Pengguna Aplikasi Jalan Kita 2.0

No	Ulasan Pengguna	Label Sentimen
1	Keren, ada jalan rusak tinggal foto	Positif
2	Tolong jangan lama, laporan sudah seminggu belum ditangani	Negatif
3	Aplikasi ini sangat membantu masyarakat	Positif

Contoh data ulasan pengguna hasil pengambilan dari Google Play Store ditunjukkan pada Tabel 1. Pengumpulan data dilakukan dengan teknik *scraping* terhadap halaman ulasan aplikasi *Jalan Kita 2.0* di Google Play Store. Proses scraping bertujuan untuk memperoleh kumpulan data ulasan dalam jumlah tertentu yang dianggap representatif. Data yang berhasil dikumpulkan kemudian disimpan dalam format terstruktur (misalnya CSV) untuk memudahkan proses pengolahan dan analisis lebih lanjut [12].

2.5. Pelabelan Data

Tabel 2. Kriteria Pelabelan Sentimen Ulasan Pengguna

Kategori Sentimen	Kriteria Umum	Contoh Ulasan
Positif	Ulasan berisi pujian, kepuasan, atau penilaian baik terhadap aplikasi atau layanan	"Aplikasinya sangat membantu dan mudah digunakan"
Negatif	Ulasan berisi keluhan, kritik, atau ketidakpuasan terhadap aplikasi atau penanganan laporan	"Laporan sudah lama tetapi belum ditangani"

Setelah data terkumpul, tahap selanjutnya adalah pelabelan sentimen. Setiap ulasan diklasifikasikan ke dalam kategori sentimen positif atau negatif berdasarkan isi teks ulasan.

Proses pelabelan dilakukan secara manual berdasarkan kriteria yang disajikan pada Tabel 2 untuk memastikan konsistensi dan mengurangi subjektivitas dalam proses klasifikasi dengan mempertimbangkan makna dan konteks kalimat untuk memastikan akurasi label. Data berlabel ini digunakan sebagai *ground truth* dalam proses pelatihan dan pengujian model klasifikasi [11].

2.6. Prapemrosesan Teks

Tabel 3. Ringkasan Data Sebelum dan Sesudah Pembersihan

Tahap Data	Jumlah Ulasan	Keterangan
Data awal hasil scraping	343	Data mentah
Setelah pembersihan	334	Data layak diolah

Jumlah data ulasan yang digunakan dalam penelitian ini mengalami pengurangan setelah proses pembersihan data, sebagaimana ditunjukkan pada Tabel 3. Prapemrosesan teks dilakukan untuk membersihkan dan menyiapkan data sebelum proses klasifikasi. Tahapan prapemrosesan meliputi:

a. *Case folding*, yaitu mengubah seluruh teks menjadi huruf kecil.
b. *Tokenizing*, yaitu memecah teks menjadi unit kata.
c. *Stopword removal*, yaitu menghapus kata-kata umum yang tidak memiliki makna signifikan dalam analisis sentimen.
d. *Stemming*, yaitu mengubah kata ke bentuk dasar.

Tahapan ini bertujuan untuk mengurangi noise dan meningkatkan kualitas fitur yang digunakan dalam proses klasifikasi [13].

Tabel 4. Contoh Hasil Prapemrosesan Teks

Teks Asli	Setelah Prapemrosesan
Keren, ada jalan rusak tinggal foto	keren jalan rusak tinggal foto
Laporannya lama diproses	lapor lama proses

Contoh perubahan teks ulasan setelah melalui tahapan tokenisasi dan stemming ditunjukkan pada Tabel 4. Hasil prapemrosesan teks menunjukkan bahwa tahapan pembersihan dan normalisasi data berperan penting dalam meningkatkan kualitas ulasan pengguna sebelum dilakukan proses klasifikasi sentimen. Pada tahap *case folding*, seluruh teks ulasan diubah menjadi huruf kecil sehingga perbedaan penulisan huruf kapital tidak memengaruhi proses analisis. Selanjutnya, proses tokenisasi memecah teks menjadi unit kata yang lebih sederhana, sehingga struktur kalimat yang kompleks dapat direpresentasikan dalam bentuk yang lebih mudah diolah. Tahap *stopword removal* berhasil menghilangkan kata-kata umum yang tidak memiliki kontribusi signifikan terhadap penentuan sentimen, seperti kata penghubung dan kata keterangan yang sering muncul. Proses stemming kemudian mengubah kata berimbuhan ke bentuk dasar, sehingga variasi kata dengan makna yang sama dapat direpresentasikan secara konsisten. Berdasarkan hasil prapemrosesan tersebut, teks ulasan menjadi lebih ringkas, bersih, dan terfokus pada kata-kata yang mencerminkan opini pengguna, sehingga mampu meningkatkan efektivitas ekstraksi fitur dan kinerja algoritma Naïve Bayes dalam mengklasifikasikan sentimen ulasan pengguna aplikasi *Jalan Kita 2.0 [14]*.

2.7. Ekstraksi Fitur

Setelah prapemrosesan, data teks direpresentasikan ke dalam bentuk numerik menggunakan metode *Term Frequency–Inverse Document Frequency (TF-IDF).* Metode ini digunakan untuk memberikan bobot pada setiap kata berdasarkan tingkat kepentingannya dalam suatu dokumen dan keseluruhan kumpulan dokumen. Representasi TF-IDF digunakan sebagai fitur masukan pada algoritma Naïve Bayes. Representasi teks ke dalam bentuk numerik dilakukan menggunakan metode TF-IDF sebagaimana ditunjukkan pada Persamaan 1, 2, dan 3.

Rumus Term Frequency (TF):

$$TF(t,d) = \frac{f_{t,d}}{\sum f_{t,d}} \quad (1)$$

Rumus Inverse Document Frequency (IDF):

$$IDF(t) = log\left(\frac{N}{df_t}\right) \quad (2)$$

Rumus TF-IDF:

$$TF\text{-}IDF(t,d) = TF(t,d) \times IDF(t) \quad (3)$$

Keterangan:

$f_{t,d}$: frekuensi kata $t$dalam dokumen d

$\sum f_{t,d}$: jumlah seluruh kata dalam dokumen d

N: jumlah seluruh dokumen

df_t: jumlah dokumen yang mengandung kata t

Metode Term Frequency–Inverse Document Frequency (TF-IDF) digunakan untuk merepresentasikan teks ulasan ke dalam bentuk numerik agar dapat diproses oleh algoritma klasifikasi. Nilai Term Frequency (TF) menunjukkan seberapa sering suatu kata muncul dalam satu dokumen ulasan, sehingga mencerminkan tingkat kepentingan kata tersebut dalam konteks dokumen tertentu. Semakin sering sebuah kata muncul dalam satu ulasan, maka nilai TF-nya semakin besar.

Sementara itu, Inverse Document Frequency (IDF) digunakan untuk mengurangi bobot kata-kata yang terlalu umum dan sering muncul di banyak dokumen, seperti kata-kata umum yang kurang informatif. IDF memberikan bobot lebih tinggi pada kata-kata yang jarang muncul namun bersifat lebih diskriminatif dalam membedakan sentimen antar dokumen. Kombinasi TF dan IDF menghasilkan bobot TF-IDF, yang merepresentasikan tingkat kepentingan suatu kata dalam sebuah dokumen relatif terhadap seluruh kumpulan dokumen. Representasi TF-IDF ini digunakan sebagai vektor fitur masukan pada algoritma Naïve Bayes untuk proses klasifikasi sentimen.

2.8. Konfigurasi dan Parameter Algoritma Naïve Bayes

Proses klasifikasi sentimen dilakukan menggunakan algoritma Multinomial Naïve Bayes, yang sesuai untuk data teks berbasis frekuensi kata. Model ini bekerja dengan menghitung probabilitas kemunculan suatu kata pada masing-masing kelas sentimen berdasarkan Teorema Bayes dengan asumsi independensi antar fitur. Dalam penelitian ini, parameter alpha (Laplace smoothing) diset dengan nilai default sebesar 1.0 untuk menghindari probabilitas nol pada kata-kata yang tidak muncul dalam data latih. Pemilihan parameter default dilakukan karena dataset yang digunakan memiliki ukuran sedang dan distribusi kata yang cukup representatif, sehingga tidak memerlukan penyesuaian parameter yang kompleks.

2.9. Pembagian Data

Dataset yang telah melalui tahap prapemrosesan dan ekstraksi fitur dibagi menjadi data latih dan data uji menggunakan skema hold-out validation dengan rasio 80:20, di mana 80% data digunakan untuk melatih model dan 20% data digunakan untuk menguji kinerja model pada data yang belum pernah dilihat sebelumnya. Skema ini dipilih untuk memperoleh gambaran performa model secara langsung dan sederhana pada data uji [15].

Meskipun teknik cross-validation belum diterapkan dalam penelitian ini, pemilihan hold-out validation dianggap memadai untuk mengevaluasi kinerja awal model klasifikasi sentimen pada dataset ulasan aplikasi Jalan Kita 2.0. Hasil evaluasi kemudian dianalisis menggunakan metrik akurasi, precision, recall, dan F1-score untuk menilai kemampuan model dalam mengklasifikasikan sentimen positif dan negatif. Penggunaan teknik validasi yang lebih kompleks, seperti k-fold cross-validation, dapat menjadi rekomendasi pengembangan pada penelitian selanjutnya guna meningkatkan reliabilitas hasil evaluasi.

2.10. Klasifikasi Menggunakan Naïve Bayes

Proses klasifikasi sentimen dilakukan menggunakan *algoritma Naïve Bayes*. Algoritma ini menghitung probabilitas kemunculan suatu kelas sentimen berdasarkan fitur kata yang terdapat dalam ulasan. Model Naïve Bayes dilatih menggunakan data latih, kemudian digunakan untuk memprediksi kelas sentimen pada data uji [16]. Proses klasifikasi sentimen dalam penelitian ini menggunakan algoritma Naïve Bayes yang didasarkan pada Teorema Bayes sebagaimana ditunjukkan pada Persamaan 4.

$$P(C_k \mid X) = \frac{P(X|C_k)\,P(C_k)}{P(X)} \qquad (4)$$

Keterangan:

$P(C_k \mid X)$= probabilitas dokumen Xtermasuk ke dalam kelas sentimen C_k

$P(X \mid C_k)$= probabilitas kemunculan fitur Xpada kelas C_k

$P(C_k)$= probabilitas awal (prior) kelas sentimen C_k

$P(X)$= probabilitas kemunculan data X

Karena $P(X)$bernilai sama untuk semua kelas, maka klasifikasi ditentukan berdasarkan nilai probabilitas terbesar $P(X \mid C_k)P(C_k)$.

2.11. Evaluasi Model

Evaluasi kinerja model dilakukan menggunakan beberapa metrik, yaitu *akurasi, precision, recall, dan F1-score*. Selain itu, *confusion matrix* digunakan untuk melihat distribusi hasil klasifikasi antara kelas sentimen aktual dan hasil prediksi. Hasil evaluasi ini digunakan untuk menilai efektivitas algoritma Naïve Bayes dalam menganalisis sentimen ulasan pengguna aplikasi *Jalan Kita 2.0*. Kinerja model dievaluasi menggunakan metrik akurasi, precision, recall, dan F1-score yang dirumuskan pada Persamaan 5, 6, 7, dan 8.

Akurasi

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

Precision

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

Recall

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

F1-score

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

Keterangan:

TP (*True Positive*)

TN (*True Negative*)

FP (*False Positive*)

FN (*False Negative*)

Metrik ini digunakan untuk menilai kinerja model klasifikasi sentimen [11].

3. HASIL PENELITIAN DAN PEMBAHASAN

3.1. Deskripsi Dataset

Dataset yang digunakan dalam penelitian ini berupa ulasan pengguna aplikasi *Jalan Kita 2.0* yang diperoleh dari Google Play Store. Setelah melalui proses pengumpulan data dan pembersihan awal, diperoleh sejumlah ulasan berbahasa Indonesia yang layak dianalisis. Ulasan-ulasan tersebut kemudian diberi label sentimen ke dalam dua kelas, yaitu positif dan negatif.

Tabel 5. Distribusi Sentimen Ulasan Pengguna

Kategori Sentimen	Jumlah Ulasan	Persentase (%)
Positif	272	81.44
Negatif	62	18.56
Total	334	100

Distribusi sentimen ulasan pengguna aplikasi Jalan Kita 2.0 berdasarkan data hasil pengolahan disajikan pada Tabel 4 Distribusi Sentimen Ulasan Pengguna. Dari total 334 ulasan yang dianalisis, sentimen positif menunjukkan proporsi paling dominan, yaitu sebanyak 272 ulasan atau sebesar 81,44%, sedangkan sentimen negatif berjumlah 62 ulasan atau sebesar 18,56%. Dominasi sentimen positif mengindikasikan bahwa mayoritas pengguna menilai aplikasi Jalan Kita 2.0 bermanfaat dalam mendukung proses pelaporan kondisi infrastruktur jalan, terutama dari aspek kemudahan penggunaan dan fungsionalitas pelaporan yang disediakan. Meskipun demikian, keberadaan sentimen negatif tetap menunjukkan adanya sejumlah kendala yang dirasakan pengguna, seperti keterlambatan penanganan laporan atau permasalahan teknis aplikasi, yang dapat dijadikan sebagai masukan penting bagi pengelola layanan untuk melakukan evaluasi dan perbaikan secara berkelanjutan. Dengan demikian, hasil distribusi sentimen pada Tabel 4 tidak hanya mencerminkan tingkat kepuasan pengguna secara umum, tetapi juga memberikan gambaran mengenai aspek layanan yang masih perlu ditingkatkan dalam rangka meningkatkan kualitas pelayanan publik berbasis aplikasi.

3.2. Hasil Prapemrosesan dan Ekstraksi Fitur

Tahap prapemrosesan teks berhasil menghilangkan unsur-unsur yang tidak relevan seperti perbedaan huruf kapital, kata umum (*stopwords*), dan variasi bentuk kata. Proses tokenisasi

dan stemming menghasilkan representasi teks yang lebih ringkas dan seragam, sehingga mengurangi noise pada data seperti tabel 4.

Tabel 6. Kata Dominan Berdasarkan Hasil TF-IDF

Sentimen	Kata Dominan
Positif	mudah, bantu, lapor, cepat, manfaat
Negatif	lama, error, login, gagal, proses

Selanjutnya, metode TF-IDF digunakan untuk merepresentasikan teks ulasan ke dalam bentuk numerik sebagaimana ditunjukkan pada Tabel 6. Hasil ekstraksi fitur menunjukkan bahwa kata-kata yang sering muncul pada ulasan positif berkaitan dengan kemudahan pelaporan dan manfaat aplikasi bagi pengguna, seperti kemudahan dalam menyampaikan laporan dan kejelasan informasi yang disediakan. Di sisi lain, ulasan negatif tetap memunculkan kata-kata yang berkaitan dengan kendala aplikasi, antara lain keterlambatan penanganan laporan, kesulitan login, dan ketidaksesuaian informasi. Temuan ini menunjukkan bahwa meskipun sebagian besar pengguna memberikan penilaian positif, masih terdapat sejumlah permasalahan yang perlu menjadi perhatian pengelola aplikasi. Hasil ini memperkuat temuan distribusi sentimen yang menunjukkan dominasi sentimen positif pada ulasan pengguna.

3.3. Hasil Klasifikasi Sentimen Menggunakan Naïve Bayes

Proses klasifikasi sentimen dilakukan dengan membagi data ke dalam data latih dan data uji menggunakan rasio tertentu. Model Naïve Bayes dilatih menggunakan data latih dan diuji menggunakan data uji untuk memprediksi kelas sentimen ulasan.

Tabel 7. Hasil Evaluasi Klasifikasi Sentimen Menggunakan Naïve Bayes

Metrik Evaluasi	Nilai
Accuracy	0.93
Precision	0.93
Recall	0.93
F1-Score	0.92

Hasil klasifikasi sentimen menggunakan algoritma Multinomial Naïve Bayes ditunjukkan pada Tabel 8. Pengujian dilakukan pada 67 data uji yang merupakan bagian dari total 334 ulasan pengguna dengan skema pembagian data 80:20. Berdasarkan hasil pengujian, model menghasilkan nilai akurasi sebesar 0,93, dengan nilai precision, recall, dan F1-score masing-masing sebesar 0,93, 0,93, dan 0,92 berdasarkan *weighted average*. Nilai evaluasi tersebut menunjukkan bahwa algoritma Naïve Bayes mampu mengklasifikasikan sentimen ulasan pengguna aplikasi *Jalan Kita 2.0* dengan tingkat kinerja yang baik dan relatif seimbang dalam mengidentifikasi sentimen positif dan negatif. Hasil ini mengindikasikan bahwa Naïve Bayes efektif digunakan sebagai metode klasifikasi pada data ulasan aplikasi yang memiliki karakteristik teks singkat dan tidak terstruktur, khususnya pada kondisi distribusi data yang didominasi oleh sentimen positif.

3.4. Evaluasi Model

Tabel 8. Confusion Matrix Klasifikasi Sentimen

Aktual \ Prediksi	Negatif	Positif
Negatif	6	5
Positif	0	56

Evaluasi kinerja model klasifikasi sentimen dilakukan menggunakan confusion matrix sebagaimana ditunjukkan pada Tabel 8. Berdasarkan tabel tersebut, model mampu mengklasifikasikan seluruh ulasan dengan sentimen positif secara tepat, yang ditunjukkan oleh tidak adanya kesalahan klasifikasi pada kelas positif. Namun, pada kelas sentimen negatif masih terdapat beberapa kesalahan klasifikasi, di mana sebagian ulasan negatif diprediksi sebagai sentimen positif. Kondisi ini menunjukkan bahwa model cenderung lebih sensitif terhadap kelas mayoritas, yaitu sentimen positif.

Perbedaan kinerja antar kelas sentimen ini dipengaruhi oleh distribusi data yang tidak seimbang, di mana jumlah ulasan positif jauh lebih banyak dibandingkan ulasan negatif. Meskipun demikian, nilai akurasi, precision, recall, dan F1-score yang diperoleh menunjukkan bahwa algoritma Naïve Bayes tetap mampu memberikan performa yang baik secara keseluruhan. Keterbatasan model dalam mengenali sentimen negatif dapat menjadi bahan pertimbangan untuk penelitian selanjutnya, misalnya dengan menyeimbangkan jumlah data atau menggunakan pendekatan klasifikasi yang lebih kompleks.

4. KESIMPULAN

Berdasarkan hasil penelitian, analisis sentimen ulasan pengguna aplikasi *Jalan Kita 2.0* menggunakan algoritma Multinomial Naïve Bayes menunjukkan bahwa mayoritas ulasan memiliki kecenderungan sentimen positif, yang mengindikasikan bahwa aplikasi tersebut dinilai memberikan manfaat bagi pengguna dalam proses pelaporan kondisi infrastruktur jalan. Penerapan tahapan prapemrosesan teks dan ekstraksi fitur menggunakan metode TF-IDF mampu menghasilkan representasi data yang efektif untuk proses klasifikasi sentimen. Hasil evaluasi model menunjukkan kinerja yang baik dengan nilai akurasi, precision, recall, dan F1-score yang tinggi, meskipun masih terdapat keterbatasan dalam mengenali sentimen negatif akibat ketidakseimbangan distribusi data. Oleh karena itu, hasil penelitian ini dapat dijadikan sebagai bahan evaluasi bagi pengelola aplikasi untuk meningkatkan kualitas layanan, serta sebagai dasar bagi penelitian selanjutnya dengan memperluas jumlah data, menyeimbangkan kelas sentimen, atau menerapkan metode klasifikasi yang lebih kompleks guna memperoleh hasil yang lebih optimal.

PERNYATAAN PENGGUNAAN ARTIFICIAL INTELLIGENCE

Penulis menggunakan ChatGPT sebagai pendukung dalam penulisan dan perbaikan kode pemrograman yang berkaitan dengan pelaksanaan penelitian. Output yang disarankan oleh AI diuji dan diperiksa kembali untuk memastikan kesesuaian dengan metode penelitian. Penulis bertanggung jawab atas pelaksanaan eksperimen, validasi hasil, dan pelaporan temuan.

KONTRIBUSI PENULIS

Muhammad Ghianza Al Ghifari: Penyusunan draf awal, perumusan atau pengembangan tujuan dan sasaran penelitian

Asep Saeppani: Supervisi, validasi metodologi, peninjauan naskah.

PENDANAAN

Penelitian ini tidak menerima hibah khusus dari lembaga pendanaan mana pun, baik sektor publik, komersial, maupun nirlaba.

KONFLIK KEPENTINGAN

Penulis menyatakan tidak terdapat konflik kepentingan terkait publikasi artikel ini.

DAFTAR PUSTAKA

[1] F. Aditiya, D. Maulana, and Edora, "The Use of Naïve Bayes Algorithm in Sentiment Analysis of Grab Application Reviews," *Jurnal Komputer, Informasi dan Teknologi*, vol. 4, no. 2, p. 10, 2024, doi: 10.53697/jkomitek.v4i2.1999.

[2] I. G. S. D. Putra and I. N. T. A. Putra, "Implementasi Metode Naïve Bayes Pada Analisis Sentimen Pengguna Aplikasi Mobile Kita Bisa," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 13, no. 2, 2025, doi: 10.23960/jitet.v13i2.6423.

[3] S. Syafrizal, M. Afdal, and R. Novita, "Analisis Sentimen Ulasan Aplikasi PLN Mobile Menggunakan Algoritma Naïve Bayes Classifier dan K-Nearest Neighbor," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 4, no. 1, pp. 10–19, 2023, doi: 10.57152/malcom.v4i1.983.

[4] E. Pasalli and A. Hasibuan, "Aplikasi Website Indeks Kepuasan Masyarakat Menggunakan Metode Extreme Programming Di Dinas Pengendalian Penduduk Dan Keluarga Berencana Kabupaten Minahasa," *ENGINTECH: Journal of Engineering and Technological Sciences*, vol. 1, no. 1, pp. 17–26, 2025, doi: 10.64924/34t3rb80.

[5] B. D. Samudera, Nurdin, and H. A. K. Aidilof, "Sentiment Analysis of User Reviews on BSI Mobile and Action Mobile Applications on the Google Play Store Using Multinomial Naive Bayes Algorithm," *International Journal of Engineering, Science and Information Technology*, vol. 4, no. 4, pp. 101–112, 2024, doi: 10.52088/ijesty.v4i4.581.

[6] A. A. Purnama and Y. R. Sipayung, "Sentiment Analysis of Public Service Using Naïve Bayes Classifier," *Journal of Information Systems and Informatics*, vol. 7, no. 3, pp. 2439–2457, 2025, doi: 10.51519/journalisi.v7i3.1207.

[7] R. Maheri, F. N. Salisah, F. Muttakin, and Megawati, "Analisis Sentimen Ulasan Aplikasi M-Paspor Menggunakan Naive Bayes Dan Support Vector Machine," *JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 10, no. 1, pp. 448–458, 2025, [Online]. Available: https://jurnal.stkippgritulungagung.ac.id/index.php/jipi

[8] D. Tribuana, U. Usman, and D. Dayanti, "Penerapan Natural Language Processing Untuk Analisis Sentimen Terhadap Layanan Publik Di Media Sosial Twitter," *Jurnal Teknologi dan Bisnis Cerdas*, vol. 1, no. 1, pp. 28–37, 2025, doi: 10.64476/jtbc.v1i1.3.

[9] F. Akmali, A. D. Riyanto, and I. Darmayanti, "Optimization Naïve Bayes Algorithm in Sentiment Analysis of Bukalapak App Reviews," *Sinkron*, vol. 9, no. 1, pp. 145–151, 2024, doi: 10.33395/sinkron.v9i1.13132.

[10] M. A. A. Maldini and A. Septi, "Analisis Sentimen Ulasan Pengguna Aplikasi Flip," *JATI(Jurnal Mahasiswa Teknik Informatika)*, vol. 9, no. 3, pp. 4098–4105, 2025.

[11] V. Oktaviani, B. Warsito, H. Yasin, R. Santoso, and Suparti, "Sentiment analysis of e-commerce application in Traveloka data review on Google Play site using Naïve Bayes classifier and association method," *J Phys Conf Ser*, vol. 1943, no. 1, 2021, doi: 10.1088/1742-6596/1943/1/012147.

[12] S. A. Helmayanti, F. Hamami, and R. Y. Fa'rifah, "Application of the TF-IDF and Naïve Bayes Algorithms for Aspect-Based Sentiment Analysis of Flip Application Reviews on the Google Play Store," *Jurnal Indonesia : Manajemen Informatika dan Komunikasi*, vol. 4, no. 3, pp. 1822–1834, 2023.

[13] E. S. J. Atmadji, Y. Wabula, H. T. Karsanti, and K. Kristopher, "The Use of Naive Bayes Classifier in Sentiment Analysis at Indonesia's Super Priority Tourism Destinations Based on User Reviews," *Jurnal Sosioteknologi*, vol. 24, no. 2, pp. 226–239, 2025, doi: 10.5614/sostek.itbj.2025.24.2.7.

[14] P. Sheridan, Z. Ahmed, and A. A. Farooque, "A Fisher's Exact Test Justification of the TF–IDF Term-Weighting Scheme," *American Statistician*, vol. 0, no. 0, pp. 1–24, 2025, doi: 10.1080/00031305.2025.2539241.

[15] V. B. Lestari and C. A. Hutagalung, "Evaluation of TF-IDF Extraction Techniques in Sentiment Analysis of Indonesian-Language Marketplaces Using SVM, Logistic Regression, and Naive Bayes," *J-KOMA Journal of Computer Science and Applications*, no. 021, pp. 22–2025, 2025, [Online]. Available: https://doi.org/10.21009/j-

[16] Y. He, G. Ou, P. Fournier-Viger, and J. Z. Huang, 'Attribute grouping-based naive Bayesian classifier', Sci. China Inf. Sci., vol. 68, no. 3, p. 132106, Mar. 2025, doi: 10.1007/s11432-022-3728-2.