

ANALISIS ULASAN APLIKASI DALAM GOOGLE PLAY STORE MENGGUNAKAN MODEL NAIVE BAYES

Chintia Cantika¹, Mayer Dani Sitompul², Andri Wijaya³

Sekolah

Email: chintiac008@gmail.com¹, mayerdanis11@gmail.com², Andri_wijaya@ukmc.ac.id³

Informasi	Abstract
Volume : 3 Nomor : 1 Bulan : Januari Tahun : 2026 E-ISSN : 3062-9624	<p><i>This study aims to analyze user sentiment toward mobile applications based on reviews collected from Google Play Store by applying the Naive Bayes classification model. User reviews represent an important source of information that reflects user experiences, satisfaction levels, and perceived application quality. However, the large volume and unstructured nature of textual reviews make manual analysis inefficient and subjective. Therefore, this research adopts a quantitative approach using text classification based on machine learning to automatically categorize user reviews into positive, negative, and neutral sentiment classes. The research process consists of data collection, text preprocessing, feature extraction, sentiment classification using Naive Bayes, and model performance evaluation. Text preprocessing includes case folding, tokenizing, stopword removal, and stemming to improve data quality. Model performance is evaluated using accuracy, precision, recall, and F1-score metrics. The results show that positive sentiment dominates user reviews, indicating that the application is generally well received by users, although negative and neutral sentiments remain present and highlight areas that require improvement. The evaluation results demonstrate that the Naive Bayes model achieves reliable performance in classifying sentiment, with balanced evaluation metrics that indicate stable classification capability. These findings confirm that Naive Bayes remains an effective and efficient method for sentiment analysis of application reviews. This study contributes theoretically to sentiment analysis research and practically provides insights that can support application developers in evaluating user feedback and improving application quality.</i></p>

Keyword: Sentiment analysis, application reviews, Google Play Store, Naive Bayes, machine learning, text classification

Abstrak

Penelitian ini bertujuan untuk menganalisis sentimen pengguna terhadap aplikasi mobile berdasarkan ulasan yang terdapat di Google Play Store dengan menggunakan model Naive Bayes. Ulasan pengguna merupakan sumber informasi penting yang mencerminkan pengalaman, tingkat kepuasan, serta penilaian pengguna terhadap kualitas aplikasi. Namun, jumlah ulasan yang besar dan bentuk data teks yang tidak terstruktur menjadikan analisis manual kurang efisien dan berpotensi subjektif. Oleh karena itu, penelitian ini menggunakan pendekatan kuantitatif melalui klasifikasi teks berbasis machine learning untuk mengelompokkan ulasan pengguna ke dalam kategori sentimen positif, negatif, dan netral. Tahapan penelitian meliputi pengumpulan data, pra-pemrosesan teks, ekstraksi fitur, klasifikasi sentimen menggunakan Naive Bayes, serta evaluasi kinerja model. Pra-pemrosesan teks dilakukan melalui tahapan case folding, tokenizing, stopword removal, dan stemming untuk meningkatkan kualitas data. Evaluasi kinerja model dilakukan menggunakan metrik akurasi, precision, recall, dan F1-score. Hasil penelitian menunjukkan bahwa sentimen positif mendominasi

ulasan pengguna, yang mengindikasikan bahwa aplikasi secara umum diterima dengan baik oleh pengguna, meskipun sentimen negatif dan netral tetap muncul sebagai bentuk variasi pengalaman pengguna. Hasil evaluasi kinerja menunjukkan bahwa model Naive Bayes memiliki performa yang cukup andal dan stabil dalam mengklasifikasikan sentimen ulasan. Temuan ini menunjukkan bahwa Naive Bayes masih relevan digunakan dalam analisis sentimen ulasan aplikasi serta dapat dimanfaatkan oleh pengembang sebagai bahan evaluasi untuk meningkatkan kualitas aplikasi.

Kata Kunci: Analisis sentimen, ulasan aplikasi, Google Play Store, Naive Bayes, machine learning, klasifikasi teks

A. PENDAHULUAN

Perkembangan aplikasi *mobile* dalam beberapa tahun terakhir menunjukkan peningkatan yang sangat signifikan seiring dengan meluasnya penggunaan perangkat pintar dalam kehidupan sehari-hari. Aplikasi *mobile* tidak lagi hanya dimanfaatkan sebagai sarana hiburan, tetapi telah berkembang menjadi bagian penting dalam berbagai aktivitas, seperti pendidikan, bisnis, komunikasi, hingga pelayanan publik [1]. Beragam kebutuhan pengguna kini dipenuhi melalui aplikasi yang dapat diakses secara cepat dan mudah. Dalam hal ini, platform distribusi aplikasi seperti Google Play Store berperan sebagai penghubung utama antara pengguna dan pengembang aplikasi. Melalui platform tersebut, pengguna tidak hanya dapat mengunduh dan menggunakan aplikasi, tetapi juga memberikan penilaian dan ulasan berdasarkan pengalaman penggunaan yang dirasakan. Ulasan pengguna menjadi salah satu indikator penting yang menggambarkan kualitas, kinerja, serta tingkat kepuasan pengguna terhadap suatu aplikasi [2].

Ulasan pengguna memiliki nilai yang tinggi karena bersifat terbuka dan berasal dari pengalaman nyata. Setiap ulasan umumnya mencerminkan penilaian subjektif pengguna terhadap fitur, kemudahan penggunaan, maupun permasalahan teknis yang ditemui selama menggunakan aplikasi. Dalam jumlah yang besar, ulasan tersebut dapat memberikan gambaran umum mengenai persepsi pengguna secara kolektif. Bagi pengembang aplikasi, ulasan dapat dimanfaatkan sebagai bahan evaluasi untuk meningkatkan kualitas aplikasi dan layanan yang diberikan. Namun demikian, semakin meningkatnya jumlah pengguna juga berdampak pada bertambahnya volume ulasan yang masuk setiap hari. Ketika ulasan telah mencapai ratusan hingga ribuan entri, proses analisis secara manual menjadi tidak efisien, membutuhkan waktu yang lama, serta berpotensi menimbulkan bias subjektif dari pihak penilai. Oleh karena itu, diperlukan pendekatan yang mampu mengolah data ulasan secara sistematis dan objektif agar informasi yang terkandung di dalamnya dapat dimanfaatkan secara optimal.

Ulasan aplikasi pada Google Play Store umumnya berbentuk teks bebas yang tidak terstruktur. Teks ulasan dapat terdiri dari kalimat pendek maupun panjang, menggunakan bahasa formal maupun informal, serta sering kali mengandung singkatan atau istilah sehari-hari. Selain itu, setiap pengguna memiliki gaya penulisan yang berbeda, sehingga menambah variasi dalam data ulasan. Karakteristik tersebut membuat proses identifikasi sentimen menjadi semakin kompleks apabila hanya mengandalkan pembacaan satu per satu. Perbedaan panjang ulasan, variasi kosakata, serta keberadaan opini yang ambigu dapat menyulitkan proses penilaian secara konsisten. Kondisi ini mendorong penggunaan pendekatan analisis sentimen berbasis *machine learning* sebagai solusi untuk mengolah data teks dalam skala besar. Melalui pendekatan ini, ulasan pengguna dapat diklasifikasikan secara otomatis ke dalam kategori sentimen tertentu, seperti positif, negatif, atau netral, sehingga memberikan gambaran umum mengenai persepsi pengguna terhadap aplikasi.

Berbagai algoritma klasifikasi telah dikembangkan dan digunakan dalam analisis sentimen, salah satunya adalah *Naive Bayes*. Algoritma ini termasuk dalam metode klasifikasi probabilistik yang bekerja dengan menghitung peluang suatu kelas berdasarkan fitur-fitur yang terdapat pada data. Meskipun *Naive Bayes* memiliki asumsi independensi antar fitur, metode ini masih banyak digunakan karena memiliki kelebihan dalam hal kesederhanaan dan efisiensi komputasi. Selain itu, *Naive Bayes* relatif mudah diimplementasikan dan mampu memberikan hasil yang cukup stabil pada berbagai jenis data teks. Beberapa penelitian menunjukkan bahwa *Naive Bayes* mampu bekerja dengan baik pada data berukuran besar, terutama ketika data telah melalui tahapan pra-pemrosesan yang sesuai, seperti pembersihan teks dan penghapusan kata yang tidak relevan (Shang, 2024; Anderson et al., 2024).

Penelitian terdahulu memperlihatkan bahwa *Naive Bayes* telah diterapkan secara luas pada berbagai bidang. Algoritma ini digunakan dalam sistem tenaga untuk penilaian keamanan, dalam bidang keamanan siber untuk mendeteksi serangan, serta dalam analisis perilaku dan prediksi popularitas konten digital [1]. Hasil penelitian tersebut menunjukkan bahwa meskipun tergolong sebagai metode yang telah lama dikembangkan, *Naive Bayes* tetap relevan dalam memproses data dunia nyata dengan karakteristik yang beragam [3]. Selain itu, beberapa studi juga menegaskan bahwa performa *Naive Bayes* pada data berbasis teks cukup konsisten dan dapat diandalkan. Konsistensi tersebut umumnya dipengaruhi oleh kualitas data serta tahapan pra-pemrosesan yang dilakukan sebelum proses klasifikasi [4].

Meskipun penerapan *Naive Bayes* dalam analisis sentimen telah banyak dilakukan, kajian yang secara khusus membahas ulasan aplikasi pada Google Play Store masih

menunjukkan beberapa keterbatasan. Banyak penelitian lebih menitikberatkan pada perbandingan performa algoritma klasifikasi tanpa memberikan penjelasan yang mendalam mengenai distribusi sentimen yang dihasilkan. Selain itu, hasil klasifikasi sentimen sering kali disajikan dalam bentuk angka semata, tanpa diikuti dengan interpretasi yang jelas mengenai maknanya bagi pengembang aplikasi. Kondisi tersebut menyebabkan hasil penelitian kurang optimal dalam mendukung proses evaluasi aplikasi dan pengambilan keputusan. Oleh karena itu, diperlukan penelitian yang tidak hanya fokus pada kinerja algoritma, tetapi juga mampu menyajikan gambaran sentimen pengguna secara lebih terstruktur dan mudah dipahami.

Berdasarkan permasalahan tersebut, penelitian ini berupaya mengisi celah penelitian dengan menganalisis sentimen ulasan aplikasi di Google Play Store menggunakan model *Naive Bayes*. Penelitian ini menitikberatkan pada pengolahan ulasan pengguna sebagai data teks, mulai dari tahap pra-pemrosesan hingga proses klasifikasi sentimen. Hasil klasifikasi disajikan dalam bentuk distribusi sentimen dan evaluasi kinerja model agar dapat memberikan gambaran yang lebih jelas mengenai persepsi pengguna terhadap aplikasi yang diteliti. Pendekatan ini diharapkan dapat menghubungkan analisis akademik dengan kebutuhan praktis pengembang aplikasi dalam mengevaluasi kualitas produk yang dikembangkan.

Penelitian ini menjadi penting karena memberikan pemahaman yang lebih sistematis mengenai sentimen pengguna berdasarkan data ulasan yang tersedia di Google Play Store. Melalui penerapan *Naive Bayes*, hasil klasifikasi sentimen dapat digunakan sebagai indikator awal dalam menilai kualitas aplikasi dan tingkat penerimaan pengguna. Manfaat penelitian ini meliputi kontribusi teoretis dalam pengayaan kajian analisis sentimen berbasis teks, khususnya yang berkaitan dengan penerapan metode klasifikasi probabilistik. Selain itu, penelitian ini juga memberikan manfaat praktis bagi pengembang aplikasi sebagai bahan evaluasi untuk meningkatkan kualitas layanan, serta dapat dijadikan referensi bagi penelitian selanjutnya yang mengkaji analisis ulasan pengguna pada platform distribusi aplikasi digital.

B. METODE PENELITIAN

2.1 Pendekatan Penelitian

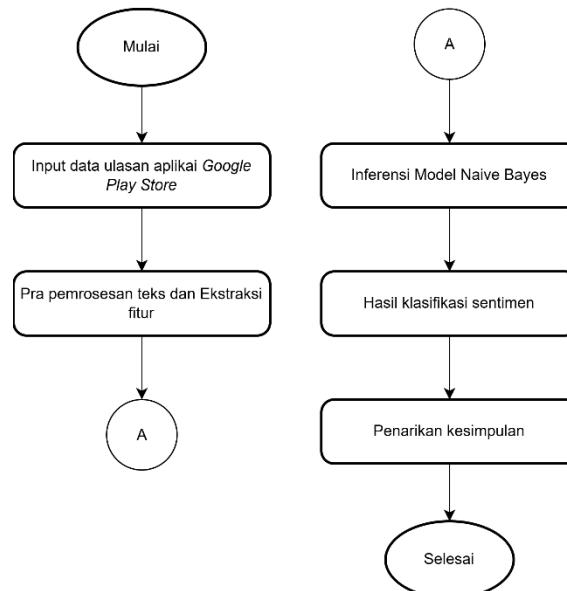
Penelitian ini menggunakan pendekatan kuantitatif dengan tujuan menganalisis sentimen ulasan pengguna aplikasi di Google Play Store secara terukur dan sistematis. Pendekatan kuantitatif dipilih karena memungkinkan pengolahan data dalam jumlah besar serta penyajian hasil dalam bentuk nilai numerik yang dapat dianalisis secara objektif.

Metode yang digunakan adalah klasifikasi teks berbasis *machine learning*, dengan fokus pada penerapan model *Naive Bayes* sebagai algoritma utama [5]. Pemilihan *Naive Bayes* didasarkan pada kemampuannya dalam menangani data teks dan efisiensinya dalam proses komputasi, sehingga sesuai untuk analisis ulasan pengguna yang bersifat tidak terstruktur. Proses penelitian dilakukan melalui tahapan pengumpulan data, pengolahan teks, penerapan model klasifikasi, serta evaluasi kinerja model untuk menilai tingkat akurasi dan keandalan hasil klasifikasi [6].

2.2 Sumber Data

Data yang digunakan dalam penelitian ini berupa ulasan pengguna aplikasi yang diperoleh dari Google Play Store. Ulasan dipilih karena mencerminkan pengalaman dan penilaian pengguna secara langsung terhadap aplikasi yang digunakan. Proses pengumpulan data dilakukan dengan mengambil ulasan pada periode waktu tertentu agar data yang diperoleh bersifat relevan dan mencerminkan kondisi penggunaan aplikasi secara aktual. Data yang dikumpulkan mencakup teks ulasan dan penilaian pengguna yang kemudian diseleksi untuk memastikan kesesuaian dengan tujuan penelitian. Ulasan yang bersifat duplikat atau tidak mengandung informasi opini pengguna dikeluarkan dari data penelitian. Seluruh data yang telah terkumpul selanjutnya digunakan sebagai bahan analisis sentimen melalui tahapan pengolahan data teks dan klasifikasi. Penggunaan data ulasan sebagai sumber penelitian telah banyak diterapkan dalam studi analisis sentimen karena mampu memberikan gambaran persepsi pengguna secara luas dan berbasis pengalaman nyata [7].

2.3 Alur Penelitian



Gambar 1. Diagram Alir Penelitian

Alur penelitian dalam studi ini disusun secara sistematis untuk menggambarkan tahapan analisis sentimen ulasan aplikasi di Google Play Store menggunakan model *Naive Bayes*. Penyusunan alur penelitian bertujuan memastikan bahwa setiap tahapan dilakukan secara terstruktur dan saling berkesinambungan, sehingga proses analisis dapat dipahami dengan jelas dan memungkinkan untuk direplikasi pada penelitian selanjutnya. Secara umum, alur penelitian dimulai dari tahap persiapan data hingga tahap penarikan kesimpulan berdasarkan hasil klasifikasi sentimen. Keseluruhan tahapan penelitian tersebut divisualisasikan dalam bentuk diagram alir penelitian yang disusun oleh peneliti dan disajikan pada gambar 1.

Tahap awal dalam alur penelitian adalah pengumpulan data ulasan pengguna aplikasi dari Google Play Store. Data yang diperoleh berupa teks ulasan yang mencerminkan pengalaman dan penilaian pengguna terhadap aplikasi yang digunakan. Data ulasan ini bersifat mentah dan memiliki tingkat variasi yang cukup tinggi, baik dari segi panjang teks, gaya bahasa, maupun penggunaan istilah informal. Kondisi tersebut menunjukkan bahwa data ulasan belum siap untuk langsung dianalisis menggunakan algoritma klasifikasi. Oleh karena itu, sebelum memasuki tahap analisis lebih lanjut, diperlukan proses pengolahan awal agar data yang digunakan memiliki kualitas yang memadai dan sesuai dengan tujuan penelitian.

Setelah data ulasan terkumpul, penelitian dilanjutkan dengan tahap pra-pemrosesan teks. Tahap ini bertujuan membersihkan dan menyiapkan data teks agar dapat diolah secara optimal oleh model klasifikasi. Pra-pemrosesan teks meliputi *case folding* untuk menyeragamkan bentuk huruf, *tokenizing* untuk memecah teks menjadi unit kata, serta *stopword removal* untuk menghilangkan kata-kata umum yang tidak memberikan kontribusi signifikan terhadap penentuan sentimen. Selain itu, dilakukan *stemming* untuk mengembalikan kata ke bentuk dasarnya sehingga variasi kata dapat dikurangi. Tahapan pra-pemrosesan ini berperan penting dalam mengurangi noise pada data teks dan meningkatkan kualitas informasi yang digunakan dalam proses analisis.

Tahap berikutnya adalah ekstraksi fitur dari teks ulasan yang telah melalui proses pra-pemrosesan. Pada tahap ini, teks ulasan direpresentasikan ke dalam bentuk fitur numerik yang dapat diproses oleh algoritma klasifikasi. Setiap ulasan dinyatakan sebagai vektor fitur yang merepresentasikan kemunculan kata-kata tertentu. Representasi fitur ini menjadi masukan utama bagi model *Naive Bayes* dalam melakukan proses klasifikasi sentimen. Kualitas fitur yang dihasilkan sangat memengaruhi hasil klasifikasi karena fitur tersebut

menjadi dasar perhitungan probabilitas pada model.

Setelah fitur diperoleh, tahap selanjutnya adalah proses inferensi menggunakan model *Naive Bayes*. Secara matematis, proses klasifikasi dilakukan dengan menentukan kelas sentimen C yang memiliki probabilitas posterior tertinggi berdasarkan fitur kata X , sebagaimana dirumuskan pada Persamaan (1):

$$P(C | X) = \frac{P(X|C) P(C)}{P(X)} \quad (1)$$

Karena nilai $P(X)$ bersifat konstan untuk seluruh kelas, maka penentuan kelas sentimen difokuskan pada nilai $P(X | C) P(C)$. Dengan asumsi independensi antar fitur, probabilitas $P(X | C)$ dapat dihitung menggunakan Persamaan (2):

$$P(X | C) = \prod_{i=1}^n P(x_i | C) \quad (2)$$

Kelas sentimen dengan nilai probabilitas tertinggi kemudian ditetapkan sebagai hasil klasifikasi, seperti positif, negatif, atau netral. Proses inferensi ini merupakan inti dari alur penelitian karena menghasilkan pengelompokan ulasan berdasarkan sentimen yang terkandung di dalamnya.

Untuk menilai keandalan hasil klasifikasi, alur penelitian dilanjutkan dengan tahap evaluasi kinerja model. Data dibagi menjadi data latih dan data uji untuk mengukur kemampuan model dalam mengklasifikasikan data yang belum pernah dipelajari sebelumnya. Evaluasi kinerja dilakukan menggunakan beberapa metrik, antara lain akurasi, *precision*, *recall*, dan *F1-score*. Akurasi dihitung menggunakan Persamaan (3):

$$\text{Akurasi} = \frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

Sementara itu, nilai *precision* dan *recall* dihitung masing-masing menggunakan Persamaan (4) dan Persamaan (5):

$$\text{Precision} = \frac{TP}{TP+FP} \quad (4)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (5)$$

Nilai *F1-score* digunakan untuk menyeimbangkan *precision* dan *recall*, sebagaimana ditunjukkan pada Persamaan (6):

$$F1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (6)$$

Tahap akhir dalam alur penelitian adalah penarikan kesimpulan berdasarkan hasil klasifikasi sentimen dan evaluasi kinerja model. Pada tahap ini, hasil analisis diinterpretasikan untuk memberikan gambaran umum mengenai persepsi pengguna

terhadap aplikasi yang diteliti. Kesimpulan yang dihasilkan tidak hanya mempertimbangkan nilai kinerja model, tetapi juga distribusi sentimen yang diperoleh dari proses klasifikasi. Dengan demikian, alur penelitian yang disusun secara sistematis ini memungkinkan penelitian menghasilkan temuan yang terstruktur, mudah dipahami, dan relevan untuk dijadikan dasar evaluasi aplikasi serta pengembangan penelitian selanjutnya.

C. HASIL DAN PEMBAHASAN

3.1 Distribusi Sentimen Ulasan Aplikasi

Distribusi sentimen ulasan aplikasi merupakan gambaran awal yang penting untuk memahami persepsi pengguna terhadap aplikasi yang dianalisis. Melalui analisis sentimen, ulasan pengguna yang semula berupa teks bebas dapat dikelompokkan ke dalam kategori sentimen tertentu, yaitu positif, negatif, dan netral. Pengelompokan ini bertujuan untuk menyederhanakan informasi yang terkandung dalam ulasan sehingga dapat diinterpretasikan secara lebih sistematis. Distribusi sentimen juga memberikan indikasi awal mengenai kecenderungan penilaian pengguna, apakah lebih dominan bersifat mendukung, mengkritisi, atau bersifat netral terhadap aplikasi yang digunakan.

Tabel 1. Data Distribusi Ulasan Aplikasi

Kategori Sentimen	Jumlah Ulasan	Persentase
Positif	1.25	55,56
Negatif	650	28,89
Netral	350	15,56
Total	2.25	100,00

Hasil distribusi sentimen ulasan aplikasi dalam penelitian ini disajikan pada Tabel 1. Berdasarkan hasil tersebut, sentimen positif merupakan kategori yang paling dominan dibandingkan dengan sentimen negatif dan netral. Dominasi sentimen positif menunjukkan bahwa sebagian besar pengguna memberikan penilaian yang baik terhadap aplikasi, baik dari segi fungsi, kemudahan penggunaan, maupun manfaat yang dirasakan [8]. Pola ini sejalan dengan karakteristik umum ulasan pada platform distribusi aplikasi, di mana pengguna cenderung lebih terdorong untuk memberikan ulasan positif ketika aplikasi mampu memenuhi kebutuhan dan ekspektasi mereka [9].

Meskipun sentimen positif mendominasi, keberadaan sentimen negatif dalam jumlah yang cukup signifikan tetap perlu mendapatkan perhatian. Sentimen negatif umumnya

berkaitan dengan keluhan pengguna terhadap masalah teknis, seperti aplikasi yang mengalami gangguan, fitur yang tidak berjalan dengan baik, atau pengalaman penggunaan yang kurang memuaskan [10]. Ulasan bersentimen negatif memberikan informasi yang bernilai karena mencerminkan aspek aplikasi yang belum mampu memenuhi harapan pengguna. Dalam hal ini, analisis distribusi sentimen tidak hanya menggambarkan tingkat kepuasan pengguna, tetapi juga berfungsi sebagai sarana untuk mengidentifikasi permasalahan yang sering dihadapi pengguna. Beberapa penelitian menyatakan bahwa ulasan negatif dapat menjadi masukan yang efektif bagi pengembang dalam meningkatkan kualitas aplikasi dan layanan yang disediakan [11].

Selain sentimen positif dan negatif, sentimen netral juga muncul sebagai bagian dari distribusi ulasan aplikasi. Ulasan bersentimen netral umumnya berisi informasi yang bersifat deskriptif tanpa menunjukkan penilaian yang jelas, seperti penyebutan pembaruan aplikasi, penjelasan singkat mengenai fitur tertentu, atau pertanyaan dari pengguna [12]. Keberadaan sentimen netral menunjukkan bahwa tidak semua ulasan mengandung unsur emosi atau evaluasi eksplisit. Dalam analisis sentimen, kategori netral memiliki peran penting karena membantu membedakan ulasan yang bersifat opini dengan ulasan yang hanya menyampaikan informasi. Dengan demikian, pembagian sentimen ke dalam tiga kategori memberikan gambaran yang lebih utuh mengenai karakteristik ulasan pengguna.

Distribusi sentimen yang diperoleh dalam penelitian ini juga mencerminkan kondisi data yang tidak sepenuhnya seimbang, di mana jumlah ulasan bersentimen positif lebih besar dibandingkan kategori lainnya. Ketidakseimbangan kelas sentimen merupakan kondisi yang umum dijumpai dalam analisis ulasan aplikasi. Kondisi ini perlu diperhatikan karena dapat memengaruhi proses dan hasil klasifikasi sentimen. Namun demikian, model *Naive Bayes* dikenal memiliki kemampuan yang cukup baik dalam menangani data dengan distribusi kelas yang tidak seimbang, khususnya pada data berbasis teks, sehingga tetap mampu menghasilkan klasifikasi yang stabil dan dapat diandalkan.

Distribusi sentimen dalam penelitian ini juga dapat digunakan untuk memberikan gambaran awal mengenai tingkat penerimaan aplikasi oleh pengguna. Dominasi sentimen positif menunjukkan bahwa aplikasi secara umum diterima dengan baik oleh pengguna. Akan tetapi, proporsi sentimen negatif yang tidak kecil mengindikasikan adanya ruang perbaikan yang masih perlu diperhatikan oleh pengembang. Informasi ini dapat dimanfaatkan untuk menentukan prioritas perbaikan, terutama pada aspek yang paling sering dikeluhkan oleh

pengguna. Dengan demikian, analisis distribusi sentimen memiliki nilai praktis dalam mendukung proses evaluasi aplikasi dan pengambilan keputusan pengembangan.

Dari sudut pandang akademik, pola distribusi sentimen yang diperoleh sejalan dengan temuan penelitian terdahulu yang menyatakan bahwa ulasan pengguna pada platform digital cenderung didominasi oleh sentimen positif, sementara sentimen negatif dan netral tetap hadir sebagai representasi pengalaman pengguna yang beragam [9]. Kesamaan pola ini menunjukkan bahwa pendekatan analisis sentimen yang digunakan dalam penelitian ini mampu menangkap kecenderungan umum dalam data ulasan aplikasi. Oleh karena itu, hasil distribusi sentimen yang diperoleh dapat dianggap representatif dalam menggambarkan persepsi pengguna terhadap aplikasi yang diteliti.

Secara keseluruhan, distribusi sentimen ulasan aplikasi memberikan dasar yang kuat untuk pembahasan lebih lanjut mengenai kinerja model klasifikasi dan implikasi hasil analisis sentimen. Informasi mengenai proporsi sentimen positif, negatif, dan netral menjadi pijakan awal dalam menilai efektivitas model *Naïve Bayes* serta memahami pola persepsi pengguna terhadap aplikasi. Pada sub-bab berikutnya, hasil distribusi sentimen ini akan dikaitkan dengan evaluasi kinerja model untuk memberikan gambaran yang lebih komprehensif mengenai hasil analisis sentimen yang dilakukan.

3.2 Evaluasi Kinerja Model *Naïve Bayes*

Evaluasi kinerja model merupakan tahap penting untuk menilai sejauh mana model *Naïve Bayes* mampu mengklasifikasikan sentimen ulasan aplikasi secara tepat dan konsisten. Evaluasi ini dilakukan setelah proses pelatihan dan pengujian model selesai, dengan tujuan memastikan bahwa hasil klasifikasi yang diperoleh tidak hanya bergantung pada data latih, tetapi juga memiliki kemampuan generalisasi terhadap data yang belum pernah dipelajari sebelumnya. Dengan demikian, evaluasi kinerja model berperan sebagai dasar dalam menilai keandalan metode yang digunakan dalam penelitian ini.

Hasil evaluasi kinerja model *Naïve Bayes* disajikan pada Tabel 2. Berdasarkan hasil tersebut, nilai akurasi yang diperoleh menunjukkan bahwa model mampu mengklasifikasikan sebagian besar ulasan ke dalam kategori sentimen yang sesuai. Akurasi yang relatif tinggi mengindikasikan bahwa kombinasi tahapan pra-pemrosesan, ekstraksi fitur, dan inferensi probabilistik yang diterapkan telah berjalan dengan baik. Nilai ini juga menunjukkan bahwa model *Naïve Bayes* cukup efektif digunakan untuk analisis sentimen ulasan aplikasi yang bersifat tidak terstruktur dan memiliki variasi bahasa yang beragam.

Tabel 2. Evaluasi Kinerja Model Naive Bayes

Metrik Evaluasi	Nilai (%)
Akurasi	82,67
Precision	81,94
Recall	80,88
F1-Score	81,41

Selain akurasi, metrik evaluasi lain yang digunakan dalam penelitian ini adalah *precision* dan *recall*. Nilai *precision* menggambarkan kemampuan model dalam menghasilkan klasifikasi sentimen yang tepat, khususnya dalam menghindari kesalahan pengelompokan ulasan ke dalam kategori sentimen tertentu. Nilai *precision* yang baik menunjukkan bahwa sebagian besar ulasan yang diprediksi ke dalam suatu kelas sentimen memang benar berasal dari kelas tersebut. Hal ini penting karena kesalahan klasifikasi, terutama pada sentimen negatif, dapat memengaruhi interpretasi hasil analisis dan pengambilan keputusan oleh pengembang aplikasi [13].

Sementara itu, nilai *recall* menunjukkan kemampuan model dalam mengenali seluruh ulasan yang benar-benar termasuk dalam suatu kelas sentimen. Nilai *recall* yang memadai menandakan bahwa model tidak terlalu banyak mengabaikan ulasan yang seharusnya terklasifikasi ke dalam kategori tertentu. Dalam analisis sentimen, keseimbangan antara *precision* dan *recall* menjadi aspek yang penting karena kedua metrik tersebut saling melengkapi. Model dengan *precision* tinggi tetapi *recall* rendah berpotensi mengabaikan banyak ulasan yang relevan, sedangkan model dengan *recall* tinggi tetapi *precision* rendah berisiko menghasilkan klasifikasi yang kurang tepat.

Untuk menyeimbangkan kedua metrik tersebut, penelitian ini juga menggunakan nilai *F1-score* sebagai indikator evaluasi. *F1-score* merupakan rata-rata harmonik dari *precision* dan *recall*, sehingga memberikan gambaran yang lebih komprehensif mengenai kinerja model. Nilai *F1-score* yang relatif seimbang menunjukkan bahwa model *Naive Bayes* tidak hanya mampu menghasilkan prediksi yang tepat, tetapi juga mampu mengenali sebagian besar ulasan yang relevan dalam setiap kategori sentimen. Hal ini memperkuat temuan bahwa model yang digunakan memiliki performa yang stabil dalam melakukan klasifikasi sentimen ulasan aplikasi.

Hasil evaluasi kinerja model dalam penelitian ini sejalan dengan temuan penelitian terdahulu yang menyatakan bahwa *Naive Bayes* memiliki performa yang kompetitif dalam klasifikasi teks, khususnya pada data ulasan atau opini pengguna. Beberapa studi

menunjukkan bahwa meskipun terdapat algoritma lain yang lebih kompleks, *Naive Bayes* tetap mampu memberikan hasil yang memadai dengan tingkat efisiensi komputasi yang tinggi [14]. Keselarasan hasil ini menunjukkan bahwa pemilihan *Naive Bayes* sebagai model klasifikasi dalam penelitian ini didukung oleh bukti empiris dari penelitian sebelumnya.

Dari sudut pandang praktis, hasil evaluasi kinerja model menunjukkan bahwa *Naive Bayes* dapat dijadikan sebagai alat awal yang efektif untuk menganalisis sentimen ulasan aplikasi. Dengan performa yang cukup baik, model ini mampu memberikan gambaran umum mengenai persepsi pengguna tanpa memerlukan sumber daya komputasi yang besar. Hal ini menjadi keuntungan tersendiri, terutama bagi pengembang aplikasi atau peneliti yang memiliki keterbatasan sumber daya. Namun demikian, hasil evaluasi juga menunjukkan bahwa masih terdapat peluang untuk meningkatkan kinerja model, misalnya melalui penambahan jumlah data, pengayaan fitur, atau penggunaan pendekatan hibrida pada penelitian selanjutnya.

Secara keseluruhan, evaluasi kinerja model *Naive Bayes* menunjukkan bahwa metode yang digunakan dalam penelitian ini mampu mengklasifikasikan sentimen ulasan aplikasi secara cukup andal. Nilai akurasi, *precision*, *recall*, dan *F1-score* yang diperoleh memberikan dasar yang kuat untuk melanjutkan pembahasan pada tahap interpretasi hasil. Pada sub-bab berikutnya, hasil evaluasi kinerja model ini akan dikaitkan dengan distribusi sentimen dan implikasinya bagi pengembang aplikasi, sehingga pembahasan yang dihasilkan menjadi lebih menyeluruh dan bermakna.

3.3 Pembahasan

Pembahasan hasil penelitian ini difokuskan pada keterkaitan antara distribusi sentimen ulasan pengguna dan kinerja model *Naive Bayes* yang telah dievaluasi pada sub-bab sebelumnya. Hasil analisis menunjukkan bahwa pola sentimen yang muncul dari ulasan pengguna dapat memberikan gambaran yang cukup jelas mengenai persepsi pengguna terhadap aplikasi. Dominasi sentimen positif mengindikasikan tingkat penerimaan aplikasi yang relatif baik, sementara keberadaan sentimen negatif dan netral mencerminkan adanya variasi pengalaman pengguna yang perlu diperhatikan secara lebih mendalam. Pola ini menegaskan bahwa ulasan pengguna merupakan sumber informasi yang bernilai dalam memahami respons pengguna secara menyeluruh.

Dari sisi metodologis, keterkaitan antara distribusi sentimen dan kinerja model menunjukkan bahwa *Naive Bayes* mampu menangkap kecenderungan umum dalam data ulasan. Meskipun data memiliki distribusi kelas yang tidak seimbang, model tetap

menunjukkan performa yang stabil. Hal ini sejalan dengan karakteristik *Naive Bayes* yang dikenal cukup efektif dalam menangani data teks dengan variasi kosakata yang tinggi. Temuan ini mendukung hasil penelitian terdahulu yang menyatakan bahwa *Naive Bayes* masih relevan digunakan dalam analisis sentimen, khususnya pada data ulasan atau opini pengguna yang bersifat tidak terstruktur [12].

Hasil evaluasi kinerja model juga memberikan dasar untuk menafsirkan keandalan klasifikasi sentimen yang dihasilkan. Nilai akurasi yang diperoleh menunjukkan bahwa sebagian besar ulasan dapat diklasifikasikan dengan tepat ke dalam kategori sentimen yang sesuai. Keseimbangan antara nilai *precision* dan *recall* menunjukkan bahwa model tidak hanya mampu menghasilkan prediksi yang tepat, tetapi juga cukup baik dalam mengenali ulasan yang relevan pada setiap kategori sentimen. Kondisi ini penting karena kesalahan klasifikasi, terutama pada sentimen negatif, dapat memengaruhi interpretasi hasil dan keputusan yang diambil berdasarkan analisis tersebut [13].

Dari sudut pandang akademik, hasil penelitian ini konsisten dengan temuan beberapa studi sebelumnya yang membahas penerapan *Naive Bayes* pada berbagai bidang klasifikasi. Penelitian terdahulu menunjukkan bahwa meskipun terdapat algoritma yang lebih kompleks, *Naive Bayes* tetap mampu memberikan performa yang kompetitif dengan tingkat efisiensi yang tinggi [14]. Keselarasan ini menunjukkan bahwa pendekatan yang digunakan dalam penelitian ini memiliki landasan empiris yang kuat dan dapat dipertanggungjawabkan secara ilmiah. Dengan demikian, hasil penelitian ini tidak berdiri sendiri, melainkan menjadi bagian dari pengembangan kajian analisis sentimen berbasis metode probabilistik.

Selain itu, pembahasan hasil penelitian juga menunjukkan bahwa distribusi sentimen dapat dimanfaatkan sebagai dasar evaluasi aplikasi dari sudut pandang pengguna. Sentimen positif yang dominan dapat diartikan sebagai indikasi bahwa fitur dan fungsi utama aplikasi telah berjalan sesuai dengan harapan pengguna. Sebaliknya, sentimen negatif memberikan sinyal adanya aspek tertentu yang masih perlu diperbaiki, seperti stabilitas aplikasi atau kemudahan penggunaan. Dalam hal ini, analisis sentimen berperan sebagai alat bantu untuk mengidentifikasi area perbaikan yang perlu diprioritaskan oleh pengembang aplikasi. Pendekatan ini sejalan dengan pandangan bahwa ulasan pengguna dapat digunakan sebagai masukan langsung dalam proses pengembangan dan peningkatan kualitas aplikasi [4].

Meskipun hasil yang diperoleh menunjukkan kinerja model yang cukup baik, penelitian ini tetap memiliki keterbatasan yang perlu diperhatikan. Salah satu keterbatasan tersebut berkaitan dengan ketergantungan model terhadap kualitas data dan tahapan pra-pemrosesan.

Variasi bahasa, penggunaan istilah informal, serta ambiguitas makna dalam ulasan dapat memengaruhi hasil klasifikasi [15]. Selain itu, penggunaan satu algoritma klasifikasi juga membatasi sudut pandang analisis, sehingga hasil yang diperoleh belum tentu mencerminkan performa terbaik yang dapat dicapai dengan pendekatan lain. Oleh karena itu, hasil penelitian ini perlu dipahami sebagai gambaran awal yang masih dapat dikembangkan lebih lanjut.

Secara keseluruhan, pembahasan hasil penelitian menunjukkan bahwa penerapan *Naive Bayes* dalam analisis sentimen ulasan aplikasi mampu memberikan informasi yang bermakna mengenai persepsi pengguna. Keterkaitan antara distribusi sentimen dan kinerja model memperlihatkan bahwa metode yang digunakan cukup efektif untuk tujuan penelitian ini. Temuan ini memberikan kontribusi dalam memperkuat pemanfaatan analisis sentimen berbasis teks sebagai pendekatan yang relevan dalam mengevaluasi aplikasi digital. Hasil pembahasan ini juga menjadi landasan yang kuat untuk penarikan kesimpulan dan penyusunan rekomendasi pada bagian selanjutnya.

D. KESIMPULAN

Bagian ini berisi kesimpulan yang menjawab hal segala permasalahan yang terdapat didalam penelitian. Banyaknya kata pada bagian ini berkisar min 350 kata.

Penelitian ini bertujuan untuk menganalisis sentimen ulasan aplikasi di Google Play Store dengan menerapkan model *Naive Bayes* sebagai pendekatan klasifikasi teks. Berdasarkan tahapan penelitian yang telah dilakukan, mulai dari pengumpulan data, prapemrosesan teks, ekstraksi fitur, hingga evaluasi kinerja model, dapat disimpulkan bahwa analisis sentimen mampu memberikan gambaran yang cukup jelas mengenai persepsi pengguna terhadap aplikasi yang diteliti. Ulasan pengguna yang semula berbentuk teks bebas dan tidak terstruktur berhasil dikelompokkan ke dalam kategori sentimen positif, negatif, dan netral, sehingga informasi yang terkandung di dalamnya menjadi lebih mudah dipahami dan dianalisis secara sistematis.

Hasil distribusi sentimen menunjukkan bahwa sentimen positif merupakan kategori yang paling dominan, yang mengindikasikan bahwa aplikasi secara umum diterima dengan baik oleh pengguna. Dominasi sentimen positif mencerminkan kepuasan pengguna terhadap fungsi, fitur, maupun manfaat aplikasi. Namun demikian, keberadaan sentimen negatif dan netral juga menunjukkan bahwa tidak semua pengguna memiliki pengalaman yang sama. Sentimen negatif memberikan indikasi adanya permasalahan atau kekurangan tertentu yang masih dirasakan oleh sebagian pengguna, sementara sentimen netral menggambarkan ulasan

yang bersifat informatif tanpa penilaian yang jelas. Dengan demikian, distribusi sentimen tidak hanya berfungsi sebagai indikator tingkat kepuasan pengguna, tetapi juga sebagai sumber informasi untuk mengidentifikasi area yang perlu diperbaiki oleh pengembang aplikasi.

Dari sisi kinerja model, hasil evaluasi menunjukkan bahwa model Naive Bayes mampu mengklasifikasikan sentimen ulasan dengan tingkat akurasi yang cukup baik. Nilai precision, recall, dan F1-score yang relatif seimbang menunjukkan bahwa model tidak hanya mampu menghasilkan prediksi yang tepat, tetapi juga cukup andal dalam mengenali ulasan yang relevan pada setiap kategori sentimen. Temuan ini sejalan dengan penelitian terdahulu yang menyatakan bahwa Naive Bayes merupakan metode yang efektif dan efisien untuk klasifikasi teks, khususnya pada data ulasan atau opini pengguna (Wahyuningsih et al., 2024; Wijaya et al., 2025). Dengan demikian, pemilihan Naive Bayes sebagai model klasifikasi dalam penelitian ini dapat dianggap tepat dan didukung oleh bukti empiris.

Meskipun demikian, penelitian ini memiliki beberapa keterbatasan yang perlu diperhatikan. Salah satu keterbatasan utama terletak pada ketergantungan model terhadap kualitas data dan tahapan pra-pemrosesan. Variasi bahasa, penggunaan istilah informal, serta ambiguitas makna dalam ulasan dapat memengaruhi hasil klasifikasi sentimen. Selain itu, penelitian ini hanya menggunakan satu algoritma klasifikasi, sehingga hasil yang diperoleh belum mencerminkan perbandingan performa dengan metode lain yang mungkin memberikan hasil yang lebih optimal. Oleh karena itu, hasil penelitian ini perlu dipahami sebagai gambaran awal yang masih dapat dikembangkan lebih lanjut.

Sebagai penutup, penelitian ini memberikan kontribusi teoretis dalam pengayaan kajian analisis sentimen berbasis teks dengan menunjukkan bahwa model Naive Bayes masih relevan dan efektif digunakan pada data ulasan aplikasi. Dari sisi praktis, hasil penelitian dapat dimanfaatkan oleh pengembang aplikasi sebagai bahan evaluasi untuk memahami persepsi pengguna dan menentukan prioritas perbaikan aplikasi. Penelitian selanjutnya disarankan untuk memperluas jumlah data, mengkaji penggunaan fitur linguistik yang lebih beragam, serta membandingkan Naive Bayes dengan algoritma klasifikasi lain agar diperoleh hasil analisis sentimen yang lebih komprehensif dan mendalam.

UCAPAN TERIMAKASIH

Penulis mengucapkan terima kasih kepada semua pihak yang telah memberikan dukungan dan kontribusi dalam penyusunan artikel ilmiah ini. Ucapan terima kasih disampaikan kepada pihak dosen dan pembimbing yang telah memberikan arahan, masukan,

serta bimbingan akademik sehingga penelitian ini dapat diselesaikan dengan baik. Selain itu, penulis juga mengapresiasi pihak-pihak yang telah menyediakan referensi ilmiah dan sumber pembelajaran yang relevan, sehingga proses penyusunan landasan teori dan metodologi penelitian dapat dilakukan secara sistematis.

Ucapan terima kasih juga ditujukan kepada rekan-rekan yang telah memberikan dukungan moral, diskusi, serta saran selama proses penelitian dan penulisan artikel ini berlangsung. Dukungan tersebut sangat membantu dalam menjaga konsistensi dan kualitas penulisan. Terakhir, penulis menyampaikan terima kasih kepada semua pihak yang secara langsung maupun tidak langsung telah berkontribusi dalam penyelesaian penelitian ini. Semoga artikel ini dapat memberikan manfaat dan menjadi referensi bagi penelitian selanjutnya di bidang analisis sentimen dan pengolahan teks.

E. DAFTAR PUSTAKA

- [1] S. Liu et al., "An online dynamic security assessment integrated scheme for power systems based on sparse multinomial naive bayes and canonical correlation forest," *Sustain. Energy, Grids Networks*, vol. 39, p. 101438, 2024, doi: <https://doi.org/10.1016/j.segan.2024.101438>.
- [2] Y. Ma, "Construction and Data Analysis of a New Media Content Popularity Prediction Model Based on Naive Bayes Algorithm," *Procedia Comput. Sci.*, vol. 261, pp. 294–302, 2025, doi: <https://doi.org/10.1016/j.procs.2025.04.207>.
- [3] K. Zhang, J. Luo, C. Zhang, Y. Qiu, M. Shen, and H. Duan, "A remote sensing-based spatial prediction framework using a Naive Bayes approach for cyanobacterial blooms in eutrophic lakes of China," *J. Hydrol. Reg. Stud.*, vol. 62, p. 102894, 2025, doi: <https://doi.org/10.1016/j.ejrh.2025.102894>.
- [4] C. J. Anderson et al., "A novel naïve Bayes approach to identifying grooming behaviors in the force-plate actometric platform," *J. Neurosci. Methods*, vol. 403, p. 110026, 2024, doi: <https://doi.org/10.1016/j.jneumeth.2023.110026>.
- [5] S. Wang, J. Ren, R. Bai, Y. Yao, and X. Jiang, "A Max-Relevance-Min-Divergence criterion for data discretization with applications on naive Bayes," *Pattern Recognit.*, vol. 149, p. 110236, 2024, doi: <https://doi.org/10.1016/j.patcog.2023.110236>.
- [6] Y. Tan, B. Sherwood, and P. P. Shenoy, "A naïve Bayes regularized logistic regression estimator for low-dimensional classification," *Int. J. Approx. Reason.*, vol. 172, p. 109239, 2024, doi: <https://doi.org/10.1016/j.ijar.2024.109239>.

- [7] J. S. Aguilar-Ruiz and C. Romero-Vargas, "XNB: A package for Class-Specific Naive-Bayes classifier," *SoftwareX*, vol. 32, p. 102347, 2025, doi: <https://doi.org/10.1016/j.softx.2025.102347>.
- [8] C. Mondal and M. J. Uddin, "Classification of short-term flood events using stochastic variable selection and Gaussian Naïve Bayes classifier: A case study of Sirajganj district, Bangladesh," *Heliyon*, vol. 11, no. 2, p. e41941, 2025, doi: <https://doi.org/10.1016/j.heliyon.2025.e41941>.
- [9] M. Romano and C. Conversano, "Stairway to heaven: An emotional journey in Divina Commedia with threshold-based Naïve Bayes classifier," *Mach. Learn. with Appl.*, vol. 19, p. 100613, 2025, doi: <https://doi.org/10.1016/j.mlwa.2024.100613>.
- [10] G.-L. Ou, Y.-L. He, P. Fournier-Viger, and J. Z. Huang, "A novel multi-source weighted naive Bayes classifier," *Inf. Sci. (Ny)*, vol. 721, p. 122568, 2025, doi: <https://doi.org/10.1016/j.ins.2025.122568>.
- [11] X. Zhou, Y. Wang, L. Zhang, A. Huang, and X. Wang, "An innovative multi-view collaborative optimization framework for Weighted Naive Bayes," *Knowledge-Based Syst.*, vol. 317, p. 113378, 2025, doi: <https://doi.org/10.1016/j.knosys.2025.113378>.
- [12] T. Wahyuningsih, D. Manongga, I. Sembiring, and S. Wijono, "Comparison of Effectiveness of Logistic Regression, Naive Bayes, and Random Forest Algorithms in Predicting Student Arguments," *Procedia Comput. Sci.*, vol. 234, pp. 349–356, 2024, doi: <https://doi.org/10.1016/j.procs.2024.03.014>.
- [13] O. Peretz, M. Koren, and O. Koren, "Naive Bayes classifier – An ensemble procedure for recall and precision enrichment," *Eng. Appl. Artif. Intell.*, vol. 136, p. 108972, 2024, doi: <https://doi.org/10.1016/j.engappai.2024.108972>.
- [14] F. K. Wijaya, T. C. Sugijono, R. Setiawan, and R. Y. Rumagit, "Comparative Analysis of Machine Learning Random Forest, Naïve Bayes and SVM for Flight Delay Classification," *Procedia Comput. Sci.*, vol. 269, pp. 1546–1555, 2025, doi: <https://doi.org/10.1016/j.procs.2025.09.096>.
- [15] Y. Shang, "Prevention and detection of DDOS attack in virtual cloud computing environment using Naive Bayes algorithm of machine learning," *Meas. Sensors*, vol. 31, p. 100991, 2024, doi: <https://doi.org/10.1016/j.measen.2023.100991>.