Jurnal umsida

cek plagiasi fix Analisis Sentimen Ulasan Aplikasi ShopeePay pada Google Play Store Menggunakan Algoritma Naïve Bayes...

bulqis

Kampus 3

Perpustakaan

Document Details

Submission ID
trn:oid:::1:3650921945

Submission Date
Sep 16, 2026, 7:36 PM GMT+7

Download Date
Sep 16, 2026, 8:27 PM GMT+7

File Name
cek_plagiasi_fix_Analisis_Sentimen_Ulasan_Aplikasi_ShopeePay_pada_Google_Play_Store_Mengg....docx

File Size
1.0 MB

20 Pages

9,764 Words

66,998 Characters

16% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Filtered from the Report

- Bibliography
- Quoted Text

Top Sources

15% Internet sources
17% Publications
17% Submitted works (Student Papers)

Integrity Flags

0 Integrity Flags for Review

No suspicious text manipulations found.

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.

Top Sources

15% Internet sources
17% Publications
17% Submitted works (Student Papers)

Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1 Student papers

Universitas Muhammadiyah Sidoarjo **15%**

2 Publication

Sri Damayanti, Husna Sari, Bobby Ardiansyah, Dicky Apdillah. "Analisis Sentimen ... **2%**

Sentiment Analysis of ShopeePay Aplication Reviews on Google Play Store Using the Naïve Bayes Algorithm with the SMOTE Technique [Analisis Sentimen Ulasan Aplikasi ShopeePay pada Google Play Store Menggunakan Algoritma Naïve Bayes dengan Teknik SMOTE]

Yumna Almas Elsaviyanto[1)], Istian Kriya Almanfaluti[*2))], Andry Rachmadany[3)], Mochamad Rizal Yulianto[4)]

[1, 2, 3, 4)]Program Studi Bisnis Digital, Universitas Muhammadiyah Sidoarjo, Indonesia
*Email Penulis Korespondensi: istian.alman@umsida.ac.id

Abstract. *Objective This study aims to analyze the sentiment of ShopeePay user reviews on the Google Play Store and to evaluate the performance of Multinomial Naive Bayes on balanced and imbalanced data distributions with the application of SMOTE. Method The data were collected through web scraping using Python and processed through cleaning, case folding, normalization, stopword removal, tokenization, stemming, and feature extraction using CountVectorizer with a Bag-of-Words approach. Two scenarios were used, each consisting of 30,032 reviews, namely a balanced dataset with 15,016 positive and 15,016 negative reviews, and an imbalanced dataset with 20,016 positive and 10,016 negative reviews. The data were split into 85% training and 15% testing data, while SMOTE was applied only to the imbalanced training data. Conclusion The balanced dataset achieved 91.03% accuracy and an F1-score of 0.91 for both classes, while the imbalanced dataset with SMOTE achieved 88.15% accuracy. These findings indicate that class distribution affects model performance, and SMOTE did not outperform data that were already balanced.*

Keywords - *Sentiment Analysis; ShopeePay; Multinomial Naïve Bayes; SMOTE; Google Play Store*

Abstrak. *Tujuan penelitian ini adalah menganalisis sentimen ulasan pengguna ShopeePay di Google Play Store serta mengevaluasi kinerja Multinomial Naive Bayes pada distribusi data seimbang dan tidak seimbang dengan penerapan SMOTE. Metode penelitian menggunakan data hasil scraping dengan Python yang diproses melalui pembersihan, case folding, normalisasi, stopword removal, tokenisasi, stemming, serta ekstraksi fitur CountVectorizer berbasis Bag-of-Words. Dua skenario digunakan, masing-masing 30.032 ulasan, yaitu balanced dataset sebanyak 15.016 ulasan positif dan 15.016 ulasan negatif serta imbalanced dataset sebanyak 20.016 ulasan positif dan 10.016 ulasan negatif. Data dibagi 85% untuk training dan 15% untuk testing, sedangkan SMOTE hanya diterapkan pada data training imbalanced. Kesimpulan menunjukkan balanced dataset menghasilkan accuracy 91,03% dan F1-score 0,91 pada kedua kelas, sedangkan imbalanced dataset dengan SMOTE menghasilkan accuracy 88,15%. Distribusi kelas terbukti memengaruhi performa model dan SMOTE belum mampu melampaui kinerja data yang telah seimbang.*

Kata Kunci - *Analisis Sentimen; ShopeePay; Multinomial Naïve Bayes; SMOTE; Google Play Store*

I. PENDAHULUAN

Pesatnya perkembangan Teknologi Informasi (TI) telah mengubah cara orang berkomunikasi, berbisnis dan memperoleh informasi. Saat ini teknologi informasi bukan sekedar alat bantu, melainkan bagian utama dari ekonomi digital. Menurut Rachmatsyah [1], teknologi berbasis media, informasi, dan infrastruktur digital memiliki peran strategis dalam memperkuat dan mengembangkan ekosistem ekonomi digital secara berkelanjutan. Salah satu inovasi besar di bidang ini adalah dompet digital (*e-wallet*), yaitu sebuah sistem pembayaran elektronik. Teknologi ini memungkinkan pengguna melakukan pembayaran tanpa uang tunai, serta menawarkan proses lebih cepat dan mudah dibandingkan metode pembayaran konvensional [2].

ShopeePay adalah layanan dompet digital yang berkembang pesat di Indonesia pada saat ini. ShopeePay merupakan layanan pembayaran digital yang terintegrasi dengan platform *e-commerce* Shopee dan banyak dimanfaatkan oleh masyarakat sebagai sarana transaksi dalam berbagai aktivitas ekonomi digital, baik secara daring maupun luring. Mawardani dan Dwijayanti menjelaskan bahwa ShopeePay menjadi salah satu dompet digital yang populer karena kemudahan penggunaan serta dukungan ekosistem *e-commerce* Shopee yang luas [3]. Layanan ini menyediakan berbagai fitur transaksi, seperti pembayaran belanja daring, *top-up* saldo, pembayaran melalui QRIS, serta integrasi dengan berbagai *merchant* mitra. Berbagai penelitian juga menunjukkan bahwa faktor seperti persepsi kemudahan penggunaan, promosi, serta tingkat kepercayaan pengguna memiliki pengaruh terhadap niat masyarakat dalam menggunakan layanan dompet digital.

Meskipun demikian, tingginya tingkat adopsi dan intensitas penggunaan ShopeePay tidak secara otomatis menjamin tingkat kepuasan pengguna yang berkelanjutan. Dalam praktiknya, pengguna masih dapat mengalami

berbagai kendala teknis maupun operasional, seperti kegagalan transaksi, keterlambatan saldo, keterbatasan fitur layanan, serta potensi permasalahan keamanan sistem. Shalsabilla menunjukkan bahwa permasalahan tersebut sering kali tercermin dalam bentuk keluhan maupun ulasan negatif yang disampaikan oleh pengguna terhadap layanan ShopeePay [4]. Oleh karena itu, memahami pengalaman dan persepsi pengguna menjadi hal yang penting bagi penyedia layanan digital dalam upaya meningkatkan kualitas layanan secara berkelanjutan.

Dalam era ekonomi digital, ulasan pengguna (*user-generated reviews*) di toko aplikasi seperti *Google Play* memiliki peran penting karena mencerminkan pengalaman nyata para pengguna. Ulasan tersebut tidah hanya membantu calon pengguna memutuskan apakah merekka perlu mencoba suatu aplikasi, tetapi juga memberikan masukan berharga bagi pengembang mengenai layanan mereka. Muzaki, Febriana, dan Cholifah menyatakan bahwa ulasan pengguna mencerminkan persepsi pelanggan terhadap kualitas layanan, yang bermanfaat dalam pengambilan keputusan untuk pengembangan sistem digital [5]. Dengan demikian, analisis sistematis ulasan pengguna merupakan kunci untuk memahami pandangan massyarakat serta mengidentifikasi aspek-aspek yang memerlukan perbaikan.

Pendekatan *sentiment analysis* merupakan salah satu cara untuk menelaah pendapat pengguna dengan memproses data teks. Maldini dan Andryana menunjukkan bahwa *sentiment analysis* terhadap ulasan aplikasi digital mampu memberikan wawasan penting mengenai pengalaman pengguna, termasuk dalam mengidentifikasi permasalahan teknis, keterbatasan fitur, serta isu keamanan yang muncul pada layanan digital [6]. Dengan memanfaatkan teknik *sentiment analysis*, perusahaan dapat memahami pola keluhan maupun apresiasi pengguna secara lebih terstruktur dan berbasis data.

Sejumlah penelitian terdahulu telah telah menggunakan berbagai metode pembelajaran *machine learning* untuk analisis sentimen pada ulasan pengguna aplikasi digital, seperti Naïve Bayes, *Support Vector Machine* (SVM) dan metode *ensemble*. Sebagai contoh Praptiwi, menggunakan *Support Vector Machine* dan *Maximum Entropy* untuk mengklasifikasikan ulasan pengguna *e-commerce* di *Google Play Store* dan menunjukkan bahwa *machine learning* efektif untuk analisis sentimen teks [7]. Namun, sebagian besar penelitian tersebut berfokus pada aplikasi *e-commerce* atau layanan *mobile banking* dengan menggunakan klasifikasi sentimen konvensional. Selain itu, beberapa penelitian membandingkan berbagai algoritma tanpa mempertimbangkan masalah distribusi kelas yang tidak seimbang pada dataset ulasan pengguna.

Dalam praktiknya, data ulasan pada platform digital seperti *Google Play Store* biasanya memiliki jumlah ulasan positif lebih banyak dibandingkan ulasan negatif. Kondisi ini dikenal sebagai permasalahan *imbalanced class*, yang dapat menimbulkan bias pada model klasifikasi dan menurunkan kemampuan model dalam mendeteksi sentimen minoritas secara akurat. Selain itu, penelitian mengenai analisis sentimen terhadap ulasan aplikasi ShopeePay masih lebih sedikit dibandingkan dengan layanan keuangan digital lainnya. Sebagian besar penelitian sebelumnya belum menggabungkan metode klasifikasi teks seperti Naïve Bayes dengan teknik penanganan data yang tidak seimbang, seperti *Synthetic Minority Oversampling Technique (SMOTE)*, dalam konteks ulasan layanan dompet digital (*e-wallet*).

Terlebih lagi, jumlah ulasan pengguna di platform digital terus meningkat dan sebagian besar berupa teks tidak terstruktur. Kondisi ini membuat analisis manual menjadi lambat dan berpotensi subjektif. Cantika, Sitompul, dan Wijaya menjelaskan bahwa volume ulasan teks bebas yang besar pada *Google Play Store* membuat proses analisis manual menjadi sulit dilakukan secara konsisten, sehingga diperlukan pendekatan analisis otomatis berbasis *machine learning* [8]. Tanpa dukungan metode analisis otomatis, perusahaan akan mengalami kesulitan dalam mengidentifikasi pola opini pengguna, memetakan permasalahan layanan, serta memantau perubahan persepsi pengguna secara berkelanjutan. Hisyam dan Ayunda juga menegaskan bahwa tanpa analisis sentimen berbasis data, masukan publik dalam jumlah besar berpotensi tidak dimanfaatkan secara optimal sebagai dasar perbaikan layanan digital [9].

Oleh karena itu, penggunaan *machine learning* untuk analisis sentimen merupakan cara yang relevan dalam memproses ulasan pengguna dalam jumlah besar. Melalui klasifikasi sentimen, ulasan dapat dikelompokkan menjadi kategori seperti positif atau negatif, sehingga memudahkan organisasi untuk memahami kepuasan pelanggan dan mengidentifikasi aspek-aspek yang perlu diperbaiki. Hal ini sejalan dengan temuan Maldini dan Andryana, yang menunjukkan bahwa analisis ulasan dapat memberikan wawasan strategis bagi manajer layanan untuk meningkatkan kualitas layanan dan pengalaman pengguna [6].

Dalam penelitian ini, analisis sentimen dilakukan menggunakan algoritma Naïve Bayes karena algoritma ini efisien dan efektif dalam menangani data teks berdimensi tinggi. Penelitian ini juga menyoroti masalah ketidakseimbangan data, di mana ulasan positif lebih dominan dibandingkan ulasan lainnya. Untuk mengatasi hal tersebut, penelitian ini menggunakan *Synthetic Minority Oversampling Technique (SMOTE)* guna meningkatkan representasi sentimen yang lebih jarang muncul dalam data pelatihan. Rahayu, Umam, dan Handayani menunjukkan bahwa penggunaan SMOTE dapat membantu menciptakan model klasifikasi yang lebih stabil pada data sentimen yang tidak seimbang [10].

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk menganalisis sentimen pada ulasan aplikasi ShopeePay di Google Play Store dengan menggunakan Naïve Bayes dan SMOTE untuk mengatasi masalah ketidakseimbangan data. Penelitian ini juga membandingkan kinerja model klasifikasi pada data yang seimbang dan

tidak seimbang untuk mengetahui sejauh mana penyeimbangan data dapat meningkatkan kualitas analisis sentimen. Penelitian ini diharapkan dapat memberikan manfaat baik secara teoretis maupun praktis. Secara teoretis, penelitian ini memberikan kontribusi terhadap pemahaman mengenai bagaimana machine learning dapat digunakan untuk menganalisis persepsi pengguna dalam layanan keuangan digital di Indonesia. Secara praktis, hasil penelitian ini diharapkan dapat membantu tim ShopeePay dalam memahami pandangan pengguna secara lebih baik serta melakukan perbaikan layanan yang lebih tepat waktu dan berbasis data nyata.

Rumusan Masalah : Tujuan penelitian ini adalah untuk menjelaskan bagaimana sentimen positif dan negatif tersebar dalam ulasan pengguna aplikasi ShopeePay di Google Play Store. Kami juga ingin menguji kinerja algoritma Naïve Bayes dalam mengklasifikasikan sentimen tersebut menggunakan metrik accuracy, precision, recall, dan F1-score. Bagian lain dari penelitian ini adalah meninjau bagaimana penggunaan Synthetic Minority Oversampling Technique (SMOTE) membantu mengatasi masalah ketidakseimbangan data serta melihat dampaknya terhadap kinerja model klasifikasi saat membandingkan dataset yang seimbang dan tidak seimbang.

Pertanyaan :

1. Bagaimana pola sentimen positif dan negatif dalam ulasan pengguna aplikasi ShopeePay di Google Play Store??
2. Seberapa baik kinerja algoritma Naïve Bayes dalam mengklasifikasikan sentimen dari ulasan tersebut berdasarkan pengukuran accuracy, precision, recall, dan F1-score?
3. Bagaimana perbandingan kinerja algoritma Naïve Bayes pada dataset yang seimbang (balanced dataset) dan dataset yang tidak seimbang (imbalanced dataset) dalam analisis sentimen ulasan ShopeePay?
4. Bagaimana pengaruh penggunaan Synthetic Minority Oversampling Technique (SMOTE) membantu meningkatkan kinerja model klasifikasi sentimen saat menangani dataset yang tidak seimbang?

Kategori SDG's : Penelitian ini berkaitan dengan Tujuan Pembangunan Berkelanjutan (SDG) 9: Industri, Inovasi, dan Infrastruktur. SDG 9 berfokus pada pembangunan infrastruktur yang tangguh, peningkatan inovasi, dan pemanfaatan teknologi digital untuk mendukung pertumbuhan ekonomi. Dalam penelitian ini, penggunaan machine learning untuk analisis sentimen pada ulasan ShopeePay di Google Play Store menunjukkan bagaimana teknologi informasi dapat mendorong inovasi dalam layanan keuangan digital. Melalui analisis ulasan pengguna, penelitian ini diharapkan dapat memberikan informasi bermanfaat yang membantu pengembang dompet digital meningkatkan layanan, memahami kebutuhan pengguna, serta membangun ekonomi digital yang lebih inovatif dan responsif terhadap kebutuhan masyarakat.

Literature Review

Analisis Sentimen dan *Text Mining*

Analisis sentimen merupakan bagian dari *Natural Language Processing* (*NLP*) yang membantu mengidentifikasi opini, sentimen, atau emosi dalam teks. Husna dan Widirahayu menjelaskan bahwa melalui pendekatan *NLP*, analisis sentimen digunakan untuk mengungkap sikap dan emosi yang diekspresikan dalam data tekstual secara sistematis dan terukur [11]. Dalam penelitian aplikasi digital, analisis sentimen banyak digunakan untuk menilai persepsi pengguna melalui ulasan yang ditulis secara spontan pada situs publik seperti Google Play Store, yang mencerminkan pandangan sebenarnya pengguna terhadap suatu aplikasi. Yoman et al. menunjukkan bahwa analisis ulasan *Google Play Store* membatu mengevaluasi persepsi dan pengalaman pengguna aplikasi digital, khususnya dalam konteks layanan keuangan berbasis aplikasi [12]. Proses ini menjadi penting karena ulasan pengguna mencerminkan pengalaman aktual pelanggan yang disampaikan secara alami tanpa dibatasi oleh struktur pertanyaan survei formal. Zahra dan Carkiman juga menemukan bahwa ulasan daring lebih autentik dan mencerminkan pengalaman pelanggan yang nyata dibandingkan instrumen survei terstruktur [13].

Sebagai bagian dari *text mining*, analisis sentimen memerlukan proses konversi data teks yang tidak terstruktur menjadi angka agar dapat diproses oleh algoritma *machine learning*. Umumnya pada proses ini dilakukan menggunakan metode seperti *bag-of-words*, *term frequency–inverse document frequency* (*TF-IDF*), serta *n-grams*. Sutriawan et al. menyatakan bahwa *TF-IDF* dan *n-grams* merupakan pendekatan yang umum digunakan dalam mengubah teks menjadi angka guna meningkatkan kinerja model klasifikasi teks [14]. Melalui representasi numerik

tersebut, algoritma klasifikasi mampu mempelajari pola linguistik yang berkaitan dengan kecenderungan sentimen positif maupun negatif secara sistematis. Temuan Fazri dan Voutama juga menunjukkan bahwa penggunaan fitur berbasis *TF-IDF* memungkinkan model klasifikasi seperti *Support Vector Machine* (*SVM*) untuk menangkap pola linguistik dan membedakan polaritas sentimen pada data teks informal di media sosial [15].

Dalam konteks ulasan aplikasi ShopeePay, teks ulasan umumnya bersifat tidak baku, banyak mengandung singkatan, serta menggunakan bahasa informal. Karakteristik tersebut menyebabkan data memiliki banyak *noise* yang dapat memengaruhi performa model klasifikasi. Oleh karena itu, langkah - langkah *preprocessing* seperti *case folding*, *cleaning*, normalisasi, *tokenization*, penghapusan *stopwords*, serta *stemming* diperlukan untuk meningkatkan kualitas data sebelum membangun model klasifikasi [16]. Kualitas terhadap *preprocessing* ini sangat menentukan kinerja model analisis sentimen, karena menentukan tingkat kejelasan informasi yang dapat dipelajari oleh algoritma klasifikasi.

Secara praktis, analisis sentimen telah terbukti efektif dalam membantu organisasi memetakan permasalahan layanan, mengevaluasi kualitas fitur, serta memahami dinamika persepsi pengguna terhadap aplikasi digital. Rahman et al. menunjukkan bahwa menganalisis ulasan pengguna pada *Google Play Store* mampu mengidentifikasi isu layanan dan aspek aplikasi yang perlu diperbaiki agar aplikasi dapat melayani pengguna dengan lebih baik dan meningkatkan kepuasan [17]. Oleh karena itu, penggunaan analisis sentimen merupakan cara yang relevan untuk memahami pandangan pengguna terhadap layanan ShopeePay.

Naïve Bayes Classifier

Naïve Bayes adalah metode klasifikasi yang menggunakan probabilitas berdasarkan *Teorema Bayes* Metode ini bekerja dengan asumsi bahwa setiap fitur dalam data bersifat independen satu sama lain. Berdasarkan asumsi tersebut, algoritma ini menghitung probabilitas suatu kelas dengan menggabungkan probabilitas dari masing-masing fiturnya. Meskipun asumsi independensi ini tidak selalu sesuai dengan situasi dunia nyata, hal tersebut menjadikan *Naïve Bayes* sebagai alat yang sederhana namun efektif untuk klasifikasi probabilistik [18].

Dalam konteks klasifikasi teks, *Naïve Bayes* dikenal cepat dan efisien serta mampu menangani data dalam jumlah yang besar. Berbagai temuan empiris menunjukkan bahwa *Naïve Bayes* memiliki kinerja yang baik dan waktu pelatihan yang singkat, sehingga sering digunakan sebagai titik awal untuk analisis sentimen pada teks tidak terstruktur [19].

Salah satu versi *Naïve Bayes* yang paling populer untuk analisis teks adalah *Multinomial Naïve Bayes* (*MNB*). Model ini menggunakan frekuensi kata dalam setiap dokumen sebagai cara untuk merepresentasikan teks secara probabilistik.. Karakteristik tersebut menjadikan *Multinomial Naïve Bayes* sangat sesuai untuk representasi teks berbasis *bag-of-words*, seperti *CountVectorizer*, di mana frekuensi kemunculan kata menjadi fitur utama dalam proses klasifikasi [20]. Penelitian juga menunjukkan bahwa MNB merupakan pilihan yang kuat untuk analisis sentimen karena kemudahan penggunaannya, efisiensi komputasinya, serta kemampuannya dalam memodelkan probabilitas kata. Studi komparatif terhadap ulasan aplikasi pada *Google Play Store* menunjukkan bahwa meskipun beberapa algoritma lain seperti *Logistic Regression* dapat mencapai tingkat akurasi yang lebih tinggi pada kondisi tertentu, *Naïve Bayes* tetap memberikan kinerja yang kompetitif dengan waktu pemrosesan yang jauh lebih cepat [21].

Selain itu, *Naïve Bayes* bekerja dengan baik bahkan pada teks yang mengandung variasi bahasa, kesalahan penulisan, maupun struktur kalimat yang tidak konsisten. Hal ini menjadikannya ideal untuk digunakan dalam analisis sentimen terhadap ulasan pengguna aplikasi digital yang umumnya beragam dan tidak terstruktur [22]. Oleh karena itu, algoritma *Naïve Bayes* dipilih sebagai metode utama untuk mengklasifikasikan sentimen dalam ulasan pengguna aplikasi ShopeePay pada penelitian ini.

Dompet Digital (*E-wallet*) dan ShopeePay

Dompet digital atau yang juga disebut *e-wallet*, merupakan jenis aplikasi yang memungkinkan pengguna menyimpan uang secara elektronik. Aplikasi ini membantu pengguna dalam melakukan pembayaran, mengirim uang, serta berbagai transaksi finansial secara lebih cepat dan efisien tanpa ketergantungan pada uang tunai atau instrumen pembayaran fisik [23]. Keberadaan *e-wallet* menjadi inovasi penting dalam sistem pembayaran modern karena mampu meningkatkan kemudahan akses transaksi serta mendukung perkembangan ekonomi digital.

Penelitian sebelumnya menunjukkan bahwa adopsi layanan *e-wallet* dipengaruhi oleh berbagai faktor utama, seperti mudah digunakan, manfaat sistem, menjamin keamanan dana, serta kepercayaan pengguna terhadap penyedia layanan. Selain itu, promosi dan insentif yang ditawarkan oleh penyedia layanan juga berperan dalam mendorong minat masyarakat untuk menggunakan layanan pembayaran digital [24].

ShopeePay merupakan salah satu layanan pembayaran digital yang terintegrasi dalam ekosistem *e-commerce* Shopee dan memiliki peran penting dalam mendukung aktivitas transaksi dalam platform tersebut. Integrasi ShopeePay dengan berbagai fitur operasional Shopee, seperti subsidi ongkos kirim, promosi, serta proses pembayaran yang cepat, berkontribusi terhadap peningkatan pengalaman pengguna dalam ekosistem digital Shopee [25]. Selain

itu, ShopeePay juga menyediakan berbagai fitur transaksi seperti pembayaran melalui QRIS, transfer dana, *cashback*, serta kemudahan proses *top-up* saldo.

Namun demikian, seperti layanan aplikasi digital lainnya, pengalaman pengguna dalam menggunakan ShopeePay sangat dipengaruhi oleh kualitas sistem dan stabilitas layanan. Permasalahan seperti keterlambatan transaksi, gangguan sistem, kesalahan saat melakukan *top-up*, maupun isu keamanan akun dapat memengaruhi persepsi pengguna terhadap kualitas layanan digital. Permasalahan teknis yang berulang terbukti dapat menurunkan tingkat kepercayaan pengguna serta memengaruhi niat penggunaan berkelanjutan terhadap layanan tersebut [26].

Dalam konteks ini, ulasan di Google Play Store dapat memberikan gambaran yang jelas mengenai kualitas sebuah aplikasi digital. Ulasan-ulasan tersebut ditulis langsung oleh pengguna dan mencerminkan pendapat mereka mengenai berbagai aspek aplikasi, seperti kemudahan penggunaan, kinerja, adanya kendala teknis, serta kualitas layanan secara keseluruhan. Informasi semacam ini dapat dimanfaatkan sebagai sumber data yang berharga dalam memahami persepsi pengguna secara lebih komprehensif [27].

Penanganan Ketidakseimbangan Data (*Imbalanced Class*)

Ketidakseimbangan distribusi kelas (*class imbalance*) merupakan permasalahan umum yang sering muncul dalam analisis sentimen, khususnya saat menangani data ulasan pengguna platform digital. Dalam banyak kasus, ulasan positif lebih banyak ditemukan dibandingkan ulasan negatif atau netral. Hal ini menyebabkan model *machine learning* lebih condong pada kelas mayoritas, sehingga model kesulitan mengidentifikasi kelas minoritas secara akurat. Firdaus et al. menunjukkan bahwa ketika sentimen positif mendominasi data, kemampuan model untuk mendeteksi kelas minoritas dapat menurun, hal ini berdampak pada nilai *recall* dan *F1-score* jika ketidakseimbangan tersebut tidak diatasi [28].

Salah satu metode yang banyak digunakan untuk mengatasi masalah ini adalah *Synthetic Minority Oversampling Technique* (*SMOTE*). Metode ini diperkenalkan oleh Chawla et al. teknik ini membantu meningkatkan representasi kelas minoritas dengan cara membuat titik data baru melalui interpolasi antara sampel minoritas yang ada dan tetangga terdekatnya dalam ruang fitur [29]. Dengan cara ini, distribusi kelas dalam dataset menjadi lebih seimbang tanpa harus melakukan duplikasi data secara acak.

Berbagai penelitian menunjukkan bahwa penerapan *SMOTE* dapat meningkatkan performa model klasifikasi pada berbagai domain, termasuk analisis sentimen pada aplikasi digital. Peningkatan tersebut umumnya tercermin dari meningkatnya nilai metrik evaluasi seperti akurasi, *precision*, *recall*, dan *F1-score*, serta berkurangnya bias model terhadap kelas mayoritas [30]. Oleh karena itu, dalam penelitian ini teknik *SMOTE* digunakan untuk menyeimbangkan proporsi kelas sentimen pada data pelatihan sehingga model klasifikasi dapat mempelajari pola sentimen secara lebih representatif.

Ulasan Pengguna dan Pengambilan Keputusan

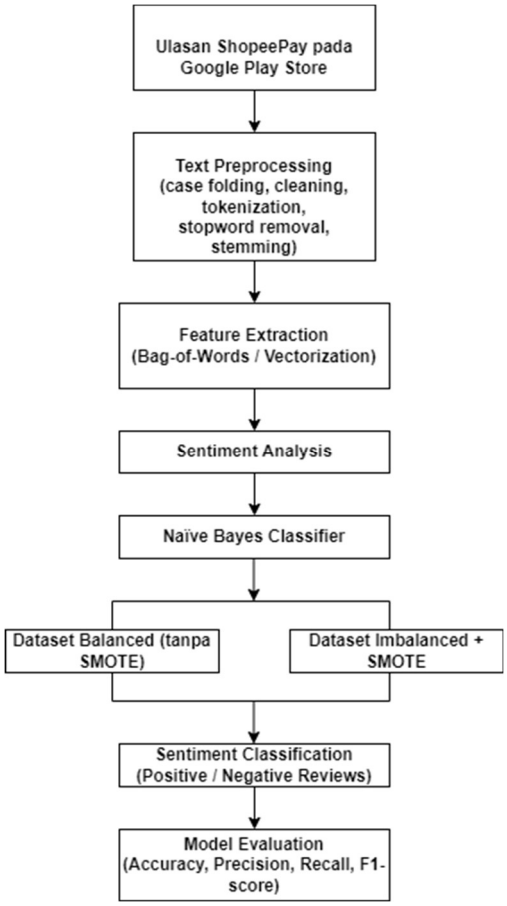
Ulasan pengguna merupakan bentuk *electronic word-of-mouth* (*e-WOM*), yaitu komunikasi elektronik yang memuat evaluasi positif maupun negatif terhadap suatu produk atau layanan yang disampaikan oleh konsumen kepada konsumen lain melalui platform digital. Penelitian menunjukkan bahwa *e-WOM* memiliki pengaruh signifikan terhadap persepsi konsumen serta keputusan penggunaan suatu produk atau layanan digital [31].

Dalam konteks aplikasi mobile, pengguna baru cenderung mengandalkan rating dan ulasan yang tersedia pada *Google Play Store* sebagai acuan utama dalam menilai kualitas dan kredibilitas suatu aplikasi sebelum melakukan instalasi. Rating dan komentar pengguna berfungsi sebagai indikator pengalaman pengguna sebelumnya yang membentuk persepsi serta tingkat kepercayaan calon pengguna. Aplikasi dengan rating tinggi dan ulasan positif cenderung memiliki tingkat unduhan yang lebih tinggi dibandingkan aplikasi dengan rating rendah dan ulasan negatif [32].

Dari perspektif penyedia layanan digital, ulasan pengguna merupakan representasi dari *voice of customer* yang memiliki nilai strategis. Ulasan tersebut memuat pengalaman nyata, evaluasi subjektif, serta ekspektasi pengguna terhadap kualitas layanan yang diterima. Informasi tersebut dapat dimanfaatkan untuk memahami persepsi pengguna, mengidentifikasi permasalahan layanan, serta merumuskan strategi peningkatan kualitas layanan secara lebih efektif [33].

Melalui pendekatan analisis sentimen berbasis data daring, perusahaan dapat mengidentifikasi isu dominan yang dihadapi pengguna, memantau perubahan persepsi pengguna dari waktu ke waktu, serta mengevaluasi dampak kebijakan maupun pembaruan fitur secara sistematis [34]. Oleh karena itu, ulasan pengguna ShopeePay pada *Google Play Store* menjadi sumber data strategis yang dapat dimanfaatkan untuk memahami tingkat kepuasan, keluhan, serta preferensi pengguna secara komprehensif, sehingga mendukung proses pengambilan keputusan berbasis data dalam upaya peningkatan kualitas layanan digital.

Kerangka Berpikir



Gambar 1. Kerangka Berpikir Penelitian
Sumber: Diolah oleh penulis (2026)

Kerangka berpikir penelitian pada gambar 1 menjelaskan bagaimana analisis sentimen terhadap ulasan pengguna aplikasi ShopeePay pada Google Play Store dilakukan. Ulasan pengguna tersebut menjadi sumber data utama karena merepresentasikan pengalaman, persepsi, serta tingkat kepuasan pengguna terhadap layanan aplikasi ShopeePay. Data ulasan yang terkumpul kemudian diproses melalui beberapa tahapan analisis yang bertujuan untuk mengklasifikasikan sentimen pengguna secara sistematis menggunakan pendekatan *machine learning*.

Tahap pertama adalah mengumpulkan data ulasan pengguna ShopeePay dari *Google Play Store*. Ulasan tersebut berbentuk data teks dan memuat berbagai opini pengguna mengenai kualitas layanan aplikasi. Selanjutnya, data tersebut diproses melalui tahap *text preprocessing* untuk membersihkan dan menyiapkan data agar dapat dianalisis lebih lanjut. Tahapan *preprocessing* meliputi *case folding* untuk mengubah semua teks menjadi huruf kecil, *cleaning* untuk menghapus karakter tambahan, *tokenization* untuk memecah teks menjadi kata-kata, *stopword removal* untuk menghilangkan kata-kata umum yang tidak memiliki makna khusus, serta *stemming* untuk mengembalikan kata ke bentuk dasarnya.

Setelah *preprocessing*, teks yang telah dibersihkan kemudian diubah menjadi angka melalui prosesyang disebut dengan tahap *feature extraction*. Proses ini dilakukan menggunakan pendekatan *bag-of-words* atau proses *vectorization* yang mengubah kata-kata menjadi vektor numerik sehingga dapat diproses oleh algoritma klasifikasi.

Tahap berikutnya adalah *sentiment analysis*, yaitu proses identifikasi dan pengelompokan sentimen dari ulasan pengguna. Dalam penelitian ini, algoritma *Naïve Bayes Classifier digunakan untuk klasifikasi sentimen*. Algoritma ini dipilih karena memiliki kecepatan kerja yang tinggi dan mampu menangani data teks dalam jumlah besar secara efektif.

Penelitian ini juga mempertimbangkan dua kondisi dataset yang berbeda, yaitu dataset yang seimbang (*balanced dataset*) dan dataset yang tidak seimbang (*imbalanced dataset*). Pada dataset yang seimbang, proporsi jumlah ulasan positif dan negatif dibuat sama sehingga proses klasifikasi dilakukan tanpa menggunakan teknik penyeimbangan data. Untuk dataset yang tidak seimbang, digunakan teknik *Synthetic Minority Oversampling Technique (SMOTE)* guna menambah jumlah ulasan yang lebih sedikit, sehingga dataset menjadi lebih seimbang.

Hasil dari kedua proses tersebut kemudian digunakan untuk melakukan *sentiment classification*, yaitu mengelompokkan ulasan pengguna ke dalam kategori sentimen positif dan negatif. Setelah klasifikasi, kinerja model dievaluasi menggunakan metrik evaluasi, yaitu *accuracy*, *precision*, *recall*, dan *F1-score*. Evaluasi ini membantu mengukur seberapa baik model dalam mengidentifikasi sentimen pengguna serta membandingkan kinerjanya pada data seimbang dan tidak seimbang, khususnya untuk melihat dampak penggunaan teknik SMOTE.

Melalui kerangka berpikir ini, penelitian diharapkan dapat memberikan gambaran mengenai distribusi sentimen pengguna ShopeePay serta mengevaluasi efektivitas penggunaan algoritma *Naïve Bayes* dan teknik *SMOTE* dalam meningkatkan kinerja model klasifikasi sentimen.

II. METODE

Penelitian ini bertujuan untuk menganalisis sentimen ulasan pengguna aplikasi ShopeePay yang diperoleh dari platform *Google Play Store*. Ulasan pengguna dipandang sebagai sumber data tekstual yang merepresentasikan pengalaman, opini, serta evaluasi pengguna terhadap kualitas layanan dompet digital. Data ulasan tersebut dianalisis menggunakan pendekatan *sentiment analysis* berbasis *machine learning*.

Secara metodologis, penelitian ini mengikuti beberapa tahapan utama, yaitu pengumpulan data ulasan, *text preprocessing*, ekstraksi fitur, proses analisis sentimen menggunakan algoritma *Naïve Bayes*, penanganan ketidakseimbangan data menggunakan teknik *Synthetic Minority Oversampling Technique* (*SMOTE*), proses klasifikasi sentimen, serta evaluasi kinerja model. Rangkaian tahapan tersebut dirancang untuk menghasilkan model klasifikasi yang mampu mengidentifikasi sentimen pengguna secara akurat dan sistematis.

Sumber Data dan Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan data sekunder berupa ulasan pengguna aplikasi ShopeePay yang diperoleh dari platform Google Play Store. Pengumpulan data dilakukan menggunakan teknik *web scraping* dengan memanfaatkan bahasa pemrograman Python untuk mengekstraksi informasi yang tersedia pada halaman ulasan aplikasi. Informasi yang dikumpulkan meliputi isi komentar, *rating* yang diberikan pengguna, tanggal publikasi ulasan, serta atribut pendukung lainnya yang diperlukan dalam proses analisis sentimen.

Proses *web scraping* menghasilkan sekitar 100.000 ulasan pengguna yang selanjutnya dijadikan sebagai data awal penelitian. Jumlah data yang besar dipilih agar mampu merepresentasikan beragam opini pengguna terhadap layanan ShopeePay sehingga model klasifikasi yang dibangun memiliki kemampuan *generalisasi* yang lebih baik.

Sebelum digunakan dalam proses analisis, seluruh data hasil *scraping* melalui tahapan *data cleaning* untuk meningkatkan kualitas *dataset*. Tahapan ini meliputi penghapusan data duplikat, penyaringan komentar yang kosong atau terlalu singkat, penghapusan karakter khusus yang tidak relevan, serta eliminasi data yang tidak memenuhi kriteria penelitian. Proses ini bertujuan menghasilkan *dataset* yang bersih, konsisten, dan siap digunakan pada tahapan pelabelan sentimen serta *text preprocessing*.

Selanjutnya, data yang telah memenuhi kriteria penelitian digunakan untuk membentuk dua skenario eksperimen, yaitu *balanced dataset* dan *imbalanced dataset*. Penyusunan kedua skenario tersebut bertujuan untuk mengevaluasi pengaruh distribusi kelas terhadap kinerja algoritma *Multinomial Naïve Bayes*, sekaligus mengukur efektivitas penerapan *Synthetic Minority Oversampling Technique* (SMOTE) dalam mengatasi permasalahan *class imbalance*.

Dengan menggunakan kedua skenario tersebut, penelitian dapat membandingkan performa model klasifikasi pada kondisi distribusi kelas yang berbeda berdasarkan metrik evaluasi, yaitu *accuracy*, *precision*, *recall*, dan *F1-score*.

Pelabelan Sentimen

Proses pelabelan sentimen dilakukan dengan memanfaatkan informasi *rating* bintang yang diberikan oleh pengguna serta isi teks ulasan. Ulasan dengan *rating* tinggi yang disertai narasi positif terhadap layanan ShopeePay dikategorikan sebagai sentimen positif. Sebaliknya, ulasan dengan *rating* rendah yang memuat keluhan atau pengalaman penggunaan yang tidak memuaskan diklasifikasikan sebagai sentimen negatif.

Ulasan yang bersifat ambigu, netral, atau mengandung kombinasi pujian dan keluhan yang sulit diinterpretasikan secara konsisten tidak dimasukkan dalam dataset penelitian. Langkah ini dilakukan untuk menjaga kejelasan batas antar kelas sentimen sehingga model klasifikasi dapat mempelajari pola sentimen secara lebih konsisten.

Text Preprocessing

Tahap *text preprocessing* bertujuan untuk menyiapkan dan membersihkan data teks sebelum digunakan dalam proses analisis. Tahapan ini diperlukan karena data ulasan pengguna umumnya memiliki struktur bahasa yang tidak seragam, sehingga perlu dilakukan penyeragaman agar data menjadi lebih efektif dan efisien untuk diproses pada tahap klasifikasi [35]. Tahapan preprocessing dalam penelitian ini meliputi case folding, cleaning, tokenization, stopword removal, dan stemming. Proses tersebut secara umum digunakan untuk mengurangi noise, menyeragamkan

bentuk teks, memisahkan teks menjadi unit kata, menghilangkan kata yang kurang informatif, serta mengembalikan kata ke bentuk dasarnya [36].

1. Case Folding
   Case folding merupakan proses mengubah seluruh huruf kapital menjadi huruf kecil sehingga teks memiliki format yang seragam [37]. Langkah ini dilakukan untuk menghindari perbedaan representasi antara kata yang memiliki makna sama tetapi ditulis menggunakan kombinasi huruf kapital dan huruf kecil yang berbeda.
2. Cleaning
   Cleaning merupakan proses menghilangkan elemen teks yang tidak diperlukan dalam analisis, seperti tanda baca, angka, URL, simbol, hashtag, dan karakter lain yang dianggap sebagai noise[36]. Tahap ini dilakukan agar data yang digunakan pada proses selanjutnya lebih bersih dan relevan.
3. Tokenization
   Tokenization merupakan proses memecah kalimat atau teks menjadi bagian yang lebih kecil berupa kata atau token sehingga setiap kata dapat diproses secara individual pada tahap analisis berikutnya [37].
4. Stopword Removal
   Stopword removal merupakan proses menghilangkan kata-kata umum yang tidak memberikan informasi penting dalam proses analisis teks sehingga fitur yang digunakan dalam klasifikasi lebih berfokus pada kata yang relevan[35].
5. Stemming
   Stemming merupakan proses mengembalikan kata berimbuhan menjadi bentuk kata dasarnya agar variasi kata yang mempunyai akar kata sama dapat diperlakukan secara konsisten dalam proses klasifikasi [36]. Dalam penelitian ini, proses stemming dilakukan menggunakan pustaka Sastrawi.

Seluruh langkah ini membantu mengurangi informasi yang tidak perlu, menyederhanakan teks, serta meningkatkan kualitas fitur yang digunakan dalam proses klasifikasi.

Ekstraksi Fitur

Setelah melalui tahapan *preprocessing*, teks ulasan diubah menjadi angka menggunakan metode *CountVectorizer* dengan pendekatan *bag-of-words*. Metode ini menghitung frekuensi kemunculan setiap kata dalam dokumen dan mengubahnya menjadi vektor numerik [20].

Pendekatan *bag-of-words* dipilih karena mampu menangkap pola distribusi kata yang berkaitan dengan sentimen dalam teks ulasan. Selain itu, metode ini sangat kompatibel dengan algoritma *Multinomial Naïve Bayes* yang dirancang untuk memodelkan data diskrit berbasis frekuensi kata.

Analisis Sentimen dengan Naïve Bayes

Langkah berikutnya melibatkan penggunaan algoritma Multinomial Naïve Bayes untuk analisis sentimen. Ini adalah metode klasifikasi yang menggunakan probabilitas berdasarkan Teorema Bayes, serta mengasumsikan bahwa setiap kata bersifat independen satu sama lain. Dalam klasifikasi teks, algoritma Multinomial Naïve Bayes menghitung probabilitas kemunculan kata dalam setiap kelas sentimen berdasarkan frekuensinya pada data pelatihan. Algoritma ini dipilih karena bekerja secara efisien dan efektif dalam menangani data teks dalam jumlah besar.

Penyeimbangan Data Menggunakan SMOTE

Dataset penelitian memiliki distribusi kelas yang tidak seimbang, dengan jumlah ulasan positif yang lebih banyak dibandingkan ulasan negatif. Hal ini dapat menyebabkan model terlalu berfokus pada kelas mayoritas dan mengabaikan kelas minoritas. Untuk mengatasinya, penelitian ini menggunakan teknik yang disebut *Synthetic Minority Oversampling Technique* (*SMOTE*). Teknik ini menambah representasi kelas minoritas dengan membentuk sampel sintetis berdasarkan interpolasi antara suatu data minoritas dan data tetangga terdekatnya dalam ruang fitur. Pengujian dilakukan melalui dua skenario. Skenario pertama menggunakan *balanced dataset* tanpa SMOTE, yaitu jumlah ulasan positif dan negatif dibuat sama. Skenario kedua menggunakan *imbalanced dataset* yang kemudian ditangani dengan SMOTE untuk menambah representasi kelas minoritas. Kedua skenario tersebut digunakan untuk membandingkan kinerja model pada kondisi distribusi data yang berbeda.

Pembangunan dan Pengujian Model

Dataset dibagi menjadi data latih (*training data*) dan data uji (*testing data*) dengan rasio 85% dan 15%. Pembagian data dilakukan secara acak menggunakan metode *Hold-Out Validation*. Data latih digunakan untuk membangun model klasifikasi sentimen, sedangkan data uji digunakan untuk mengevaluasi kemampuan model dalam mengklasifikasikan data yang belum pernah dilihat sebelumnya. Untuk menjaga validitas evaluasi model, penerapan teknik *SMOTE* hanya dilakukan pada data latih. Langkah ini dilakukan untuk mencegah terjadinya *data leakage*, sehingga data uji tetap merepresentasikan distribusi data asli.

Klasifikasi Sentimen

Setelah proses pelatihan selesai, model digunakan untuk memprediksi sentimen pada data uji. Hasil prediksi dikelompokkan ke dalam dua kelas, yaitu sentimen positif dan sentimen negatif.Klasifikasi tersebut digunakan untuk melihat kecenderungan opini pengguna terhadap layanan ShopeePay sekaligus menilai kemampuan model dalam membedakan ulasan positif dan negatif berdasarkan pola teks yang telah dipelajari.

Evaluasi Kinerja Model

Kinerja model diukur menggunakan empat metrik utama empat metrik utama, yaitu *accuracy, precision, recall*, dan *F1-score*. *Accuracy* menunjukkan jumlah prediksi yang benar dari keseluruhan prediksi. *Precision* menunjukkan seberapa baik model dalam memprediksi kelas tertentu secara tepat., sedangkan *recall* menunjukkan seberapa baik model menemukan seluruh kejadian aktual dari suatu kelas. Sementara itu, *F1-score* digunakan untuk melihat keseimbangan antara *precision* dan *recall* melalui perhitungan rata-ratanya.

Evaluasi juga mencakup *confusion matrix* dan *classification report*. Kedua hasil membantu meninjau kinerja model pada setiap kelas secara mendalam, sehingga memberikan gambaran menyeluruh mengenai kemampuannya dalam mengklasifikasikan sentimen positif dan negatif.

III. HASIL DAN PEMBAHASAN

Deskripsi Dataset Penelitian

Tabel 1. Dataset Ulasan Pengguna ShopeePay

No	Content	Score	Label
1	ok sya sudah bisa menggunakan yh dengan lancar...	5	Positif
2	terbaik dan sangat membantu	5	Positif
3	harga di shoopeepay sangat terjangkau	5	Positif
4	kalo kamu tidak mau berhenti maksa orang lain...	1	Negatif
5	Laporan terkait topup tapcash yg tidak bertambah...	5	Positif
...	…	…	...
99.996	Sangat membantu untuk transaksi 😂	5	Positif
99.997	Aplikasi yg sangat membantu	5	Positif
99.998	Hadir pengguna ke 100	5	Positif
99.999	Aplikasi yang bagus dan recommended	5	Positif
100.000	Mantaap gan, sangat membantu!	5	Positif

Sumber: Diolah oleh penulis (2026)

Berdasarkan Tabel 1, dataset awal penelitian terdiri atas sekitar 100.000 ulasan pengguna ShopeePay yang diperoleh dari Google Play Store melalui proses *web scraping*. Setiap data memuat tiga informasi utama, yaitu *content* yang berisi teks ulasan pengguna, *score* yang menunjukkan rating yang diberikan, serta *label* yang menunjukkan kategori sentimen positif atau negatif. Contoh data pada Tabel 1 memperlihatkan bahwa ulasan dengan penilaian tinggi umumnya dikategorikan sebagai sentimen positif, sedangkan ulasan yang memuat keluhan dan penilaian rendah dikategorikan sebagai sentimen negatif. Dataset tersebut selanjutnya melalui proses *data cleaning*, pelabelan, dan seleksi untuk memastikan bahwa data yang digunakan dalam tahap pemodelan memiliki kualitas yang sesuai dengan kebutuhan analisis. Dengan demikian, Tabel 1 memberikan gambaran mengenai struktur dan karakteristik awal data yang menjadi dasar dalam proses analisis sentimen pada penelitian ini.

Setelah dataset siap digunakan, penelitian membangun dua skenario, yaitu *balanced dataset* dan *imbalanced dataset*, untuk mengevaluasi pengaruh distribusi kelas terhadap performa algoritma *Multinomial Naïve Bayes*. Kedua skenario tersebut digunakan untuk membandingkan kemampuan model dalam kondisi distribusi kelas yang berbeda sekaligus mengevaluasi efektivitas penerapan *Synthetic Minority Oversampling Technique* (SMOTE) dalam menangani permasalahan *class imbalance*.

Tabel 2. Distribusi Data Berdasarkan Skenario Dataset

Skenario Dataset	Sentimen Positif	Sentimen Negatif	Total
Balanced dataset	15.016	15.016	30.032
Imbalanced dataset	20.016	10.016	30.032

Sumber: Diolah oleh penulis (2026)

Berdasarkan Tabel 2, *balanced dataset* memiliki jumlah data sentimen positif dan negatif yang sama, yaitu masing-masing sebanyak 15.016 ulasan. Kondisi ini menghasilkan distribusi kelas yang seimbang sehingga setiap kelas memiliki proporsi yang sama dalam proses pelatihan model. Distribusi yang seimbang diharapkan mampu meminimalkan kecenderungan model untuk lebih dominan mempelajari salah satu kelas sehingga hasil klasifikasi dapat dievaluasi secara lebih objektif.

Sementara itu, *imbalanced dataset* terdiri atas 20.016 ulasan positif dan 10.016 ulasan negatif. Komposisi tersebut menggambarkan kondisi yang lebih mendekati karakteristik data ulasan pada Google Play Store, di mana jumlah ulasan positif lebih banyak dibandingkan ulasan negatif. Ketidak seimbangan ini dapat menyebabkan model klasifikasi terlalu berfokus pada kelas mayoritas, sehingga kinerjanya dalam mengenali kelas minoritas menjadi kurang efektif.

Untuk mengatasi masalah *imbalance dataset* ini, digunakan teknik *Synthetic Minority Oversampling Technique (SMOTE).* Teknik ini menghasilkan sampel data baru untuk kelas minoritas guna menyeimbangkan jumlah datanya. Tujuannya adalah untuk membantu algoritma *Multinomial Naive Bayes* agar dapat lebih memahami karakteristik setiap kelas. Selanjutnya, kinerja model dibandingkan sebelum dan sesudah penggunaan SMOTE menggunakan metrik seperti *accuracy, precision, recall,* dan *F1-score* untuk melihat sejauh mana SMOTE menngkatkan hasil klasifikasi sentimen.

Preprocessing Data

Sebelum dilakukan proses klasifikasi sentimen, seluruh data ulasan pengguna yang telah melalui tahap data cleaning diproses menggunakan tahapan *text preprocessing*. Langkah ini bertujuan untuk menghasilkan data teks yang lebih sederhana, lebih seragam dan lebih sesuai untuk diproses oleh *machine learning models*. Selain itu, preprocessing dilakukan untuk mengurangi *noise* data, yang membantu menghasilkan representasi fitur yang lebih baik serta meningkatkan kinerja model.

Tahapan pertama adalah *case folding*, yaitu proses mengubah semua huruf dalam teks menjadi huruf kecil. Hal ini membantu sistem memperlakukan kata-kata yang memiliki makna sama namu ditulis dengan kombinasi huruf kapital dan huruf kecil yang berbeda sebagai kata yang sama. Sebagai contoh pada Tabel 3, kata-kata seperti *ShopeePay*, *shopeepay*, dan *SHOPEEPAY* semuanya akan berubah menjadi *shopeepay*.

Tabel 3. Hasil Case Folding

No	Input	Output (Case Folding)
1	Aplikasi Payah... Top Up Lancar Tapi Dana Gak Masuk	aplikasi payah... top up lancar tapi dana gak masuk
2	Dulu Buat Transaksi Bagus,Lancar Dr Nominal Terkecil	dulu buat transaksi bagus,lancar dr nominal terkecil
3	Sy Rubah Jd Bintang 1 Karena Tiap Buka Aplikasi...	sy rubah jd bintang 1 karena tiap buka aplikasi...
4	Aku Suka Karna Dapat DM Gratis Awok Awok DM FF	aku suka karna dapat dm gratis awok awok dm ff
5	Padahal Jarak Tempat Sortir Paket Kerumah Cuma...	padahal jarak tempat sortir paket kerumah cuma...

Sumber: Diolah oleh penulis (2026)

Setelah tahapan *case folding*, dilakukan normalisasi teks. Pada tahap ini, kata-kata tidak baku, singkatan dan istilah asing yang umum ditemukan dalam ulasan pengguna diubah menjadi bentuk baku yang tepat menggunakan kamus normalisasi khusus. Dilihat pada Tabel 4, kata *app* diubah menjadi *aplikasi*, *dr* menjadi *dari*, *gk* menjadi *tidak*, *sy* menjadi *saya*, *thanks* menjadi *terima kasih*, serta beberapa variasi penulisan lain yang memiliki makna serupa. Proses normalisasi bertujuan untuk mengurangi variasi penulisan kata sehingga informasi yang dipelajari model menjadi lebih konsisten.

Tabel 4. Hasil Normalisasi Teks

No	Input	Output (Normalisasi)
1	aplikasi payah... top up lancar tapi dana gak masuk	aplikasi payah... top up lancar tapi dana tidak masuk
2	dulu buat transaksi bagus,lancar dr nominal terkecil	dulu buat transaksi bagus,lancar dari nominal terkecil

No	Input	Output (Normalisasi)
3	sy rubah jd bintang 1 karena tiap buka aplikasi	saya rubah jadi bintang 1 karena tiap buka aplikasi
4	aku suka karna dapat dm gratis awok awok dm ff	aku suka karna dapat dm gratis awok awok dm ff
5	app shopeepay sangat membantu transaksi	aplikasi shopeepay sangat membantu transaksi

Sumber: Diolah oleh penulis (2026)

Langkah berikutnya adalah *stopword removal*, yaitu menghapus kata-kata umum yang tidak terlalu membantu dalam menentukan sentimen. Proses ini dilakukan menggunakan pustaka *Sastrawi StopWordRemover*, yang menghapus kata-kata seperti dan, yang, di, ke, serta kata-kata umum lainnya. Setelah tahap ini, teks yang tersisa sebagian besar terdiri dari kata-kata yang berkaitan dengan informasi sentimen.

Tabel 5. Hasil Stopword Removal

No	Input	Output (Stopword Removal)
1	aplikasi payah... top up lancar tapi dana tidak bisa digunakan...	aplikasi payah... top up lancar dana bisa digunakan...
2	dulu buat transaksi bagus,lancar dari nominal terkecil...	dulu buat transaksi bagus,lancar nominal terkecil...
3	saya rubah jadi bintang 1 karena tiap buka aplikasi keluar...	rubah jadi bintang 1 tiap buka aplikasi keluar...
4	aku suka karna dapat dm gratis awok awok dm ff	aku suka karna dm gratis awok awok dm ff
5	padahal jarak tempat sortir paket ke rumah cuma...	padahal jarak tempat sortir paket kerumah cuma...

Sumber: Diolah oleh penulis (2026)

Berdasarkan Tabel 5, proses *stopword removal* berhasil mengurangi keberadaan kata-kata umum yang dianggap tidak memberikan kontribusi utama dalam menentukan sentimen. Sebagai contoh, pada ulasan "aplikasi payah... top up lancar tapi dana tidak bisa digunakan..." kata *tapi* dihilangkan, sedangkan pada ulasan "dulu buat transaksi bagus, lancar dari nominal terkecil..." kata *dari* dihapus. Proses serupa juga terjadi pada kata-kata seperti *saya*, *karena*, *dapat*, dan *ke*. Dengan demikian, teks hasil *stopword removal* menjadi lebih ringkas dan lebih berfokus pada kata-kata yang memiliki informasi penting bagi proses klasifikasi sentimen. Tahap ini membantu mengurangi jumlah fitur yang kurang relevan sehingga representasi teks menjadi lebih efektif untuk diproses pada tahap berikutnya.

Selanjutnya dilakukan tokenisasi, yaitu memecah setiap kalimat menjadi kumpulan kata (*token*) menggunakan metode pemisahan berdasarkan spasi. Setiap token kemudian menjadi satuan dasar yang akan diproses pada tahap berikutnya.

Tabel 6. Hasil Tokenisasi

No	Input	Output (Tokenization)
1	aplikasi payah... top up lancar dana bisa digunakan...	[aplikasi, payah..., top, up, lancar, dana, bisa, digunakan...]
2	dulu buat transaksi bagus,lancar nominal terkecil...	[dulu, buat, transaksi, bagus,lancar, nominal, terkecil...]
3	rubah jadi bintang 1 tiap buka aplikasi keluar...	[rubah, jadi, bintang, 1, tiap, buka, aplikasi, keluar...]
4	aku suka karna dm gratis awok awok dm ff	[aku, suka, karna, dm, gratis, awok, awok, dm, ff]
5	padahal jarak tempat sortir paket kerumah cuma...	[padahal, jarak, tempat, sortir, paket, kerumah, cuma...]

Sumber: Diolah oleh penulis (2026)

Berdasarkan Tabel 6, proses tokenisasi mengubah setiap ulasan yang sebelumnya masih berbentuk kalimat menjadi kumpulan kata atau *token* secara terpisah. Sebagai contoh, teks "aplikasi payah... top up lancar dana bisa digunakan..." diubah menjadi kumpulan token berupa "[aplikasi, payah..., top, up, lancar, dana, bisa, digunakan...]". Proses yang sama diterapkan pada seluruh ulasan sehingga setiap kata dapat diperlakukan sebagai unit analisis tersendiri. Hasil tokenisasi ini menjadi dasar bagi tahapan pemrosesan berikutnya karena model membutuhkan representasi kata yang lebih terstruktur sebelum dilakukan stemming dan ekstraksi fitur. Dengan demikian, tokenisasi membantu mempersiapkan data teks agar dapat diolah secara sistematis oleh algoritma klasifikasi.

Tahapan terakhir disebut *stemming*, dimana setiap kata diubah menjadi bentuk dasarnya menggunakan pustaka *Sastrawi Stemmer*. Berdasarkan tabel 7 sebagai contoh, kata-kata *membayar*, *dibayar*, dan *pembayaran* semuanya akan diubah menjadi kata dasar *bayar*. Proses stemming bertujuan untuk menyamakan berbagai bentuk kata, sehingga fitur yang diperoleh lebih mudah diolah dan merepresentasikan makna dengan lebih baik.

Tabel 7. Hasil Stemming

No	Input (Setelah Tokenization)	Output (Stemming)
1	[aplikasi, payah..., top, up, lancar, dana, bisa, digunakan...]	aplikasi payah top up lancar dana bisa guna...
2	[dulu, buat, transaksi, bagus,lancar, nominal, terkecil...]	dulu buat transaksi bagus lancar nominal kecil...
3	[rubah, jadi, bintang, 1, tiap, buka, aplikasi, keluar...]	rubah jadi bintang 1 tiap buka aplikasi keluar...
4	[aku, suka, karna, dm, gratis, awok, awok, dm, ff]	aku suka karna dm gratis awok awok dm ff
5	[padahal, jarak, tempat, sortir, paket, kerumah, cuma, 3...]	padahal jarak tempat sortir paket rumah cuma 3...

Sumber: Diolah oleh penulis (2026)

Setelah seluruh tahapan preprocessing selesai dilakukan, hasil stemming disimpan sebagai dataset teks yang telah bersih dan siap digunakan pada tahap ekstraksi fitur. Selanjutnya, data teks tersebut dikonversi ke dalam bentuk numerik menggunakan metode *CountVectorizer (Bag of Words)* sebelum digunakan sebagai masukan pada proses pelatihan dan pengujian model Multinomial Naïve Bayes.

Pengolahan pada *Balanced Dataset*

Setelah seluruh tahapan *preprocessing* selesai dilakukan, data hasil pengolahan digunakan untuk membentuk balanced dataset. Pada skenario ini jumlah data sentimen positif dan negatif dibuat sama sehingga kedua kelas memiliki proporsi yang seimbang, yaitu masing-masing sebanyak 15.016 ulasan, dengan total 30.032 data. Keseimbangan jumlah data tersebut bertujuan untuk memberikan kesempatan yang sama bagi algoritma dalam mempelajari karakteristik masing-masing kelas tanpa dipengaruhi dominasi salah satu kelas. Dengan demikian, model diharapkan mampu menghasilkan performa klasifikasi yang lebih objektif dan mengurangi potensi bias terhadap kelas tertentu.

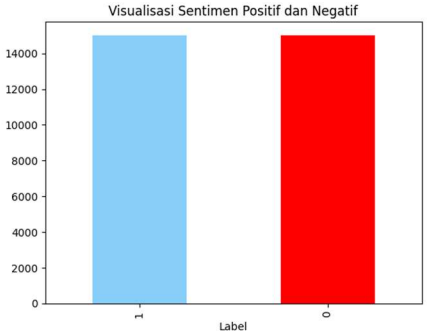
Selanjutnya, dataset dibagi menjadi 85% data latih (*training data*) dan 15% data uji (*testing data*) menggunakan metode *hold-out validation*. Pembagian dilakukan secara acak (*random state*) yang tetap sehingga proses eksperimen dapat direproduksi pada penelitian selanjutnya. Data latih yang diperoleh kemudian dikonversi menjadi representasi numerik menggunakan metode CountVectorizer dengan pendekatan Bag of Words. Metode ini mengubah data teks menjadi matriks frekuensi kata sehingga setiap dokumen direpresentasikan sebagai sekumpulan fitur numerik berdasarkan frekuensi kemunculan setiap kata.

Setelah teks diubah menjadi angka, data tersebut digunakan untuk melatih algoritma Multinomial Naïve Bayes. Algoritma ini mengidentifikasi seberapa sering setiap kata muncul dalam kaitannya dengan berbagai kategori sentimen, yang membantu membangun model untuk memprediksi sentimen ulasan baru. Setelah model dilatih, model tersebut diuji menggunakan data uji. Hasil dari model dibandingkan dengan label sebenarnya untuk mengevaluasi kinerjanya. Untuk mengukur kualitas model, kami menggunakan *confusion matrix, accuracy, precision, recall,* dan *F1-score* guna mengevaluasi kemampuannya dalam mengklasifikasikan sentimen ulasan pengguna ShopeePay secara tepat.

Visualisasi Data

Visualisasi data dilakukan sebagai tahap awal analisis untuk memperoleh gambaran mengenai karakteristik dataset sebelum dilakukan proses pelatihan model klasifikasi. Tahap ini bertujuan untuk mengetahui distribusi kelas sentimen serta mengidentifikasi kata-kata yang paling dominan muncul pada masing-masing kelas menggunakan WordCloud. Visualisasi tidak hanya memberikan informasi mengenai keseimbangan dataset, tetapi juga membantu memahami pola linguistik yang terkandung dalam ulasan pengguna ShopeePay.

Distribusi Sentimen Dataset *Balance*



Gambar 2. Visualisasi Distribusi Sentimen pada Balanced Dataset
Sumber: Diolah oleh penulis (2026)

Berdasarkan Gambar 2, dataset yang digunakan merupakan balanced dataset yang terdiri atas 30.032 ulasan, dengan komposisi 15.016 ulasan positif dan 15.016 ulasan negatif. Komposisi tersebut menunjukkan bahwa kedua kelas memiliki jumlah data yang sama sehingga tidak terdapat dominasi salah satu kelas sentimen.

Keseimbangan jumlah data merupakan salah satu faktor penting dalam proses klasifikasi karena memungkinkan algoritma Multinomial Naïve Bayes mempelajari karakteristik kedua kelas secara proporsional. Pada kondisi dataset yang tidak seimbang (*imbalanced dataset*), model cenderung membangun probabilitas yang lebih besar terhadap kelas mayoritas sehingga performa klasifikasi pada kelas minoritas dapat menurun. Sebaliknya, pada balanced dataset probabilitas awal (*prior probability*) setiap kelas menjadi relatif sama sehingga model lebih berfokus pada pola kemunculan kata dibandingkan jumlah data pada masing-masing kelas.

Hasil visualisasi ini menunjukkan bahwa proses pembentukan balanced dataset telah berhasil menghasilkan distribusi kelas yang ideal. Kondisi tersebut diharapkan mampu meningkatkan kemampuan model dalam mengenali sentimen positif maupun negatif secara lebih seimbang sehingga hasil klasifikasi menjadi lebih akurat dan tidak bias terhadap salah satu kelas.

WordCloud Sentimen Positif

WordCloud sentimen positif digunakan untuk menampilkan kata-kata yang paling sering muncul pada ulasan pengguna yang memiliki label positif. Ukuran setiap kata pada WordCloud menunjukkan frekuensi kemunculan kata tersebut dalam dataset. Semakin besar ukuran kata, semakin tinggi frekuensi kemunculannya.

Gambar 3. Word Cloud Sentimen Positif pada Balanced Dataset
Sumber: Diolah oleh penulis (2026)

Berdasarkan Gambar 3, kata-kata seperti sangat membantu, mantap, mudah, belanja, sangat baik, belanja dan shopeepay memiliki ukuran yang relatif lebih besar dibandingkan kata lainnya. Hal tersebut menunjukkan bahwa sebagian besar pengguna memberikan apresiasi terhadap kemudahan penggunaan aplikasi, kecepatan proses transaksi, serta berbagai promo yang ditawarkan oleh ShopeePay.

Dominasi kata-kata bernilai positif mengindikasikan bahwa mayoritas pengguna merasa puas terhadap layanan yang diberikan. Selain itu, visualisasi ini menunjukkan bahwa proses preprocessing telah berhasil menghilangkan

kata-kata yang tidak memiliki makna penting sehingga kata-kata yang muncul lebih mampu merepresentasikan opini pengguna secara nyata.

WordCloud Sentimen Negatif

WordCloud sentimen negatif digunakan untuk mengetahui kata-kata yang paling sering muncul pada ulasan yang mengandung keluhan pengguna terhadap aplikasi ShopeePay.

Gambar 4. Word Cloud Sentimen Negatif pada Balanced Dataset
Sumber: Diolah oleh penulis (2026)

Berdasarkan Gambar 4, kata-kata seperti aplikasi, tidak, lama, kecewa dan shopeepay muncul dengan frekuensi yang cukup tinggi. Kemunculan kata-kata tersebut menunjukkan bahwa sebagian besar keluhan pengguna berkaitan dengan kegagalan transaksi, keterlambatan saldo masuk, maupun kendala teknis yang terjadi pada aplikasi.

Informasi yang diperoleh melalui WordCloud memberikan gambaran mengenai aspek layanan yang masih memerlukan perbaikan. Dengan mengetahui kata-kata yang paling dominan pada ulasan negatif, pengembang aplikasi dapat menentukan prioritas pengembangan sistem sehingga kualitas layanan dapat terus ditingkatkan.

Pembagian Data Training dan Testing

Tabel 8. Pembagian Data Training dan Testing pada Balanced Dataset

Dataset	Jumlah Data	Persentase
Data Training	25.527	85%
Data Testing	4.505	15%
Total	**30.032**	**100%**

Sumber: Diolah oleh penulis (2026)

Berdasarkan Tabel 8, balanced dataset yang berjumlah 30.032 ulasan dibagi menjadi 25.527 data training atau 85% dan 4.505 data testing atau 15%. Komposisi tersebut menunjukkan bahwa sebagian besar data digunakan untuk membangun dan melatih model Multinomial Naïve Bayes, sedangkan sebagian lainnya digunakan untuk menguji kemampuan model dalam mengklasifikasikan data yang belum pernah digunakan pada tahap pelatihan. Pembagian data dengan proporsi 85:15 memberikan porsi data pelatihan yang cukup besar sehingga model dapat mempelajari pola kata pada masing-masing kelas sentimen secara lebih optimal, sementara data testing tetap tersedia dalam jumlah yang memadai untuk menghasilkan evaluasi kinerja yang objektif. Dengan demikian, pembagian data pada Tabel 8 menjadi dasar penting dalam proses pembangunan dan pengujian model pada skenario balanced dataset.

Ekstraksi Fitur Menggunakan *CountVectorizer*

Sebelum dilakukan proses klasifikasi, data teks hasil preprocessing dikonversi menjadi bentuk numerik menggunakan CountVectorizer dengan pendekatan Bag of Words. CountVectorizer membangun kosakata (*vocabulary*) dari seluruh kata unik yang terdapat pada data training, kemudian menghitung frekuensi kemunculan setiap kata pada masing-masing dokumen.

Melalui proses tersebut, setiap ulasan direpresentasikan sebagai sebuah vektor numerik yang menggambarkan jumlah kemunculan kata-kata tertentu dalam dokumen. Representasi numerik inilah yang digunakan sebagai masukan

pada proses pelatihan model. Penggunaan CountVectorizer dipilih karena sesuai dengan karakteristik algoritma Multinomial Naïve Bayes yang memanfaatkan distribusi frekuensi kata dalam menghitung probabilitas setiap kelas sentimen.

Pembangunan Model Multinomial Naïve Bayes

Setelah proses ekstraksi fitur selesai dilakukan, tahap berikutnya adalah membangun model klasifikasi menggunakan algoritma Multinomial Naïve Bayes. Model dilatih menggunakan 25.527 data training yang telah dikonversi ke dalam bentuk representasi numerik.

Pada tahap pelatihan, algoritma menghitung probabilitas kemunculan setiap kata terhadap masing-masing kelas sentimen berdasarkan Teorema Bayes. Hasil perhitungan tersebut membentuk model probabilistik yang selanjutnya digunakan untuk memprediksi kelas sentimen pada data testing. Multinomial Naïve Bayes dipilih karena memiliki proses komputasi yang sederhana, efisien, serta mampu memberikan performa yang baik pada permasalahan klasifikasi teks dengan jumlah data yang besar.

Evaluasi Model

Tabel 9. Hasil Evaluasi Model Multinomial Naïve Bayes pada Balanced Dataset

Kelas	Precision	Recall	F1-Score	Support
Negatif	0,90	0,92	0,91	2.254
Positif	0,92	0,90	0,91	2.251
Accuracy			0,91	4.505
Macro Average	0,91	0,91	0,91	4.505
Weighted Average	0,91	0,91	0,91	4.505

Sumber: Diolah oleh penulis (2026)

Berdasarkan Tabel 9, model Multinomial Naïve Bayes pada balanced dataset menunjukkan performa klasifikasi yang relatif seimbang pada kedua kelas sentimen. Pada kelas negatif, model memperoleh precision sebesar 0,90, recall sebesar 0,92, dan F1-score sebesar 0,91 dengan support sebanyak 2.254 data. Sementara itu, pada kelas positif diperoleh precision sebesar 0,92, recall sebesar 0,90, dan F1-score sebesar 0,91 dengan support sebanyak 2.251 data. Nilai macro average sebesar 0,91 untuk precision, recall, dan F1-score menunjukkan bahwa performa model pada kedua kelas berada pada tingkat yang relatif setara. Nilai weighted average yang juga sebesar 0,91 mengindikasikan bahwa kinerja model tetap konsisten ketika mempertimbangkan jumlah data pada masing-masing kelas.

Secara keseluruhan, nilai akurasi sebesar 91,03% menunjukkan bahwa model mampu mengklasifikasikan sekitar sembilan dari setiap sepuluh ulasan dengan benar. Nilai precision, recall, dan F1-score yang relatif seimbang pada kedua kelas menunjukkan bahwa model memiliki kemampuan yang baik dalam mengenali sentimen positif maupun negatif tanpa menunjukkan kecenderungan yang kuat terhadap salah satu kelas. Hasil ini menunjukkan bahwa penggunaan balanced dataset mampu mendukung proses pembelajaran model yang lebih proporsional sehingga performa klasifikasi menjadi lebih stabil dan akurat.

Pengolahan pada Imbalanced Dataset

Setelah seluruh tahapan *preprocessing* selesai dilakukan, data hasil pengolahan digunakan sebagai imbalanced dataset. Berbeda dengan skenario balanced dataset, pada skenario ini distribusi jumlah data antara sentimen positif dan sentimen negatif dibiarkan sesuai kondisi asli hasil pengumpulan data tanpa dilakukan proses penyeimbangan kelas. Dataset yang digunakan terdiri atas 30.032 ulasan, yang meliputi 20.016 ulasan positif dan 10.016 ulasan negatif. Komposisi tersebut menunjukkan bahwa jumlah data sentimen positif hampir dua kali lebih banyak dibandingkan data sentimen negatif sehingga membentuk kondisi class imbalance.

Distribusi kelas yang tidak seimbang merupakan salah satu permasalahan yang umum dijumpai pada penelitian analisis sentimen. Ketidakseimbangan tersebut dapat memengaruhi proses pembelajaran algoritma karena model cenderung lebih banyak mempelajari karakteristik kelas mayoritas dibandingkan kelas minoritas. Akibatnya, model memiliki kecenderungan menghasilkan prediksi yang lebih akurat pada kelas mayoritas, sedangkan kemampuan dalam mengenali kelas minoritas menjadi kurang optimal.

Selanjutnya, dataset dibagi menjadi 85% data training dan 15% data testing menggunakan metode *hold-out validation*. Pembagian dilakukan secara acak (*random state*) yang tetap sehingga hasil eksperimen dapat direproduksi pada penelitian selanjutnya. Berdasarkan proses tersebut diperoleh 25.527 data training dan 4.505 data testing.

Berbeda dengan skenario balanced dataset, pada penelitian ini data training yang masih memiliki distribusi tidak seimbang diproses menggunakan metode Synthetic Minority Oversampling Technique (SMOTE). Penerapan SMOTE dilakukan setelah proses train-test split sehingga hanya data training yang mengalami proses oversampling, sedangkan data testing tetap mempertahankan distribusi aslinya. Langkah ini bertujuan untuk menghindari terjadinya data leakage, yaitu kondisi ketika informasi dari data uji secara tidak langsung digunakan selama proses pelatihan model sehingga dapat menyebabkan hasil evaluasi menjadi bias.

SMOTE bekerja dengan menghasilkan data sintetis pada kelas minoritas berdasarkan karakteristik data yang telah ada, bukan hanya menduplikasi data secara acak. Proses ini dilakukan dengan membentuk sampel baru melalui perhitungan jarak terhadap beberapa tetangga terdekat (*k-nearest neighbors*) pada kelas minoritas. Dengan demikian, distribusi data training menjadi lebih seimbang sehingga algoritma memperoleh jumlah contoh yang relatif sama pada setiap kelas selama proses pembelajaran.

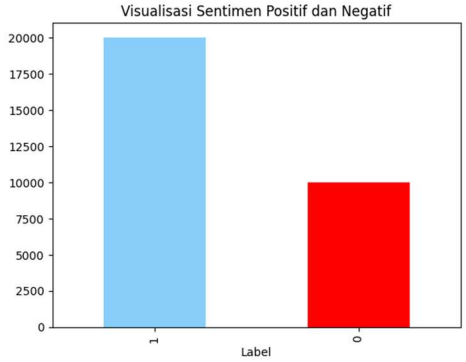
Penggunaan SMOTE pada penelitian ini diharapkan mampu meningkatkan kemampuan algoritma Multinomial Naïve Bayes dalam mengenali pola sentimen negatif yang sebelumnya memiliki jumlah data lebih sedikit. Setelah proses oversampling selesai dilakukan, data training hasil SMOTE dikonversi ke dalam bentuk representasi numerik menggunakan CountVectorizer dengan pendekatan Bag of Words. Representasi numerik tersebut selanjutnya digunakan sebagai masukan pada proses pelatihan algoritma Multinomial Naïve Bayes.

Model yang telah dibangun kemudian diuji menggunakan 4.505 data testing yang tetap menggunakan distribusi asli tanpa proses oversampling. Hasil prediksi dibandingkan dengan label sebenarnya untuk memperoleh nilai accuracy, precision, recall, F1-score, dan confusion matrix sebagai dasar evaluasi performa model. Seluruh metrik evaluasi tersebut selanjutnya digunakan untuk menganalisis efektivitas penerapan SMOTE dalam meningkatkan kemampuan model mengklasifikasikan sentimen pada dataset yang memiliki distribusi kelas tidak seimbang.

Visualisasi Data

Visualisasi data dilakukan untuk memberikan gambaran mengenai distribusi kelas sentimen sebelum dan sesudah penerapan *Synthetic Minority Oversampling Technique (SMOTE)*. Tahap ini bertujuan untuk mengetahui perubahan proporsi kelas setelah dilakukan proses oversampling serta mengidentifikasi kata-kata yang paling dominan muncul pada masing-masing kelas sentimen menggunakan *WordCloud*. Visualisasi tersebut juga digunakan untuk menunjukkan bahwa SMOTE berhasil mengurangi permasalahan *class imbalance* pada data latih sebelum proses pelatihan model dilakukan.

Distribusi Sentimen Sebelum SMOTE



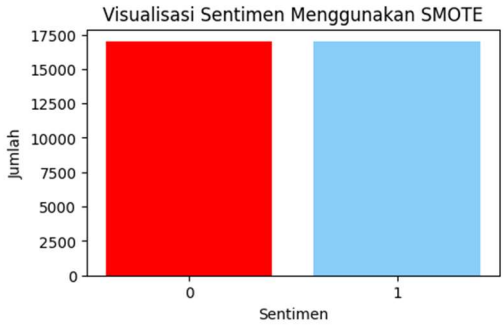
Gambar 5. Visualisasi Distribusi Sentimen Sebelum Penerapan SMOTE
Sumber: Diolah oleh penulis (2026)

Berdasarkan Gambar 5, dataset hasil preprocessing terdiri atas 30.032 ulasan, yang meliputi 20.016 ulasan positif dan 10.016 ulasan negatif. Distribusi tersebut menunjukkan bahwa jumlah data sentimen positif jauh lebih banyak dibandingkan sentimen negatif sehingga membentuk kondisi *imbalanced dataset*.

Ketidakseimbangan distribusi data tersebut berpotensi memengaruhi proses pembelajaran algoritma Multinomial Naïve Bayes karena model memperoleh lebih banyak contoh dari kelas mayoritas dibandingkan kelas minoritas. Kondisi tersebut dapat menyebabkan model memiliki kecenderungan menghasilkan prediksi pada kelas mayoritas sehingga kemampuan dalam mengenali kelas minoritas menjadi kurang optimal.

Distribusi Sentimen Setelah SMOTE

Setelah dilakukan proses *Synthetic Minority Oversampling Technique (SMOTE)* pada data training, distribusi kelas menjadi lebih seimbang. SMOTE menghasilkan data sintetis pada kelas minoritas sehingga jumlah data pada kedua kelas memiliki proporsi yang sama.



Gambar 6. Visualisasi Distribusi Sentimen Setelah Penerapan SMOTE
Sumber: Diolah oleh penulis (2026)

Visualisasi pada Gambar 6 menunjukkan bahwa proses oversampling berhasil mengurangi ketimpangan distribusi kelas tanpa mengubah data testing. Dengan distribusi data yang lebih seimbang, algoritma memperoleh jumlah contoh yang relatif sama pada setiap kelas sehingga proses pembelajaran menjadi lebih optimal.

Penerapan SMOTE hanya dilakukan pada data training, sedangkan data testing tetap menggunakan distribusi asli. Strategi tersebut bertujuan untuk menghindari *data leakage*, sehingga hasil evaluasi model tetap mencerminkan kemampuan model dalam menghadapi data nyata yang memiliki distribusi tidak seimbang.

WordCloud Sentimen Positif

Gambar 7. Word Cloud Sentimen Positif pada Imbalanced Dataset
Sumber: Diolah oleh penulis (2026)

WordCloud sentimen positif menunjukkan kata-kata yang paling sering muncul pada ulasan pengguna dengan label positif. Semakin besar ukuran suatu kata, semakin tinggi frekuensi kemunculannya dalam dataset. Berdasarkan Gambar 7, kata-kata seperti sangat membantu, mantap, bagus, aplikasi, dan shopeepay memiliki ukuran yang relatif lebih besar dibandingkan kata lainnya. Hal tersebut menunjukkan bahwa sebagian besar pengguna memberikan tanggapan positif terhadap kemudahan penggunaan aplikasi, kecepatan transaksi, serta berbagai program promosi yang ditawarkan oleh ShopeePay.

Dominasi kata-kata tersebut mengindikasikan bahwa mayoritas ulasan pengguna berisi pengalaman positif dalam menggunakan layanan pembayaran digital ShopeePay. Selain itu, hasil WordCloud juga menunjukkan bahwa tahapan preprocessing telah berhasil menghilangkan kata-kata yang tidak memiliki informasi penting sehingga kata-kata yang muncul lebih mampu merepresentasikan opini pengguna.

WordCloud Sentimen Negatif

Gambar 8. Word Cloud Sentimen Negatif pada Imbalanced Dataset
Sumber: Diolah oleh penulis (2026)

WordCloud sentimen negatif memperlihatkan kata-kata yang paling sering muncul pada ulasan yang mengandung keluhan pengguna terhadap aplikasi ShopeePay. Berdasarkan Gambar 8, kata-kata seperti aplikasi, shopeepay, tidak, transaksi dan saldo muncul dengan ukuran yang relatif besar. Kata-kata tersebut menunjukkan bahwa sebagian besar ulasan negatif berkaitan dengan kegagalan transaksi, keterlambatan saldo masuk, proses verifikasi akun, maupun kendala teknis lainnya.

Informasi tersebut memberikan gambaran mengenai aspek layanan yang masih perlu diperbaiki oleh pengembang aplikasi. Selain itu, WordCloud juga menunjukkan bahwa meskipun jumlah data negatif lebih sedikit dibandingkan data positif, ulasan negatif tetap mengandung informasi yang penting untuk meningkatkan kualitas layanan.

Pembagian Data Training dan Testing

Tabel 10. Pembagian Data Training dan Testing pada Imbalanced Dataset

Dataset	Jumlah Data	Persentase
Data Training	25.527	85%
Data Testing	4.505	15%
Total	**30.032**	**100%**

Sumber: Diolah oleh penulis (2026)

Berdasarkan Tabel 10, *imbalanced* dataset yang berjumlah 30.032 ulasan dibagi menjadi 25.527 data *training* atau 85% dan 4.505 data *testing* atau 15%. Pembagian tersebut menunjukkan bahwa sebagian besar data digunakan pada tahap pelatihan model, sedangkan sisanya digunakan untuk menguji kemampuan model dalam mengklasifikasikan ulasan yang belum pernah dipelajari sebelumnya. Pada skenario *imbalanced dataset*, data training selanjutnya menjadi bagian yang diproses dalam penerapan SMOTE, sedangkan data testing tetap digunakan sebagai dasar evaluasi performa model. Dengan komposisi 85:15, model memperoleh jumlah data pelatihan yang memadai untuk mempelajari karakteristik kata dan pola sentimen, sementara jumlah data testing tetap cukup untuk menilai kemampuan generalisasi model secara objektif. Dengan demikian, pembagian data pada Tabel 10 menjadi tahapan penting sebelum proses penyeimbangan kelas dan evaluasi model pada skenario *imbalanced dataset*.

Ekstraksi Fitur Menggunakan CountVectorizer

Sebelum dilakukan proses klasifikasi, seluruh data teks hasil preprocessing diubah menjadi representasi numerik menggunakan *CountVectorizer* dengan pendekatan *Bag of Words*. *CountVectorizer* membangun kosakata (*vocabulary*) dari seluruh kata unik yang terdapat pada data training, kemudian menghitung frekuensi kemunculan setiap kata pada masing-masing dokumen.

Melalui proses tersebut, setiap ulasan direpresentasikan sebagai sebuah vektor numerik yang menggambarkan jumlah kemunculan kata tertentu dalam dokumen. Representasi numerik tersebut selanjutnya digunakan sebagai masukan pada proses pelatihan model Multinomial Naïve Bayes.

Pembangunan Model Multinomial Naïve Bayes

Model klasifikasi dibangun menggunakan algoritma Multinomial Naïve Bayes karena algoritma ini memiliki kemampuan yang baik dalam mengolah data teks berbasis frekuensi kata. Model dilatih menggunakan 25.527 data training yang telah dikonversi menjadi bentuk numerik menggunakan *CountVectorizer*.

Pada tahap pelatihan, algoritma menghitung probabilitas kemunculan setiap kata terhadap masing-masing kelas sentimen berdasarkan Teorema Bayes. Probabilitas tersebut kemudian digunakan sebagai dasar dalam menentukan kelas sentimen dari setiap ulasan pada data testing.

Evaluasi Model

Tabel 11. Hasil Evaluasi Model Multinomial Naïve Bayes pada Imbalanced Dataset

Kelas	Precision	Recall	F1-Score	Support
Negatif	0,76	0,94	0,84	1.510
Positif	0,96	0,85	0,91	2.995
Accuracy			0,88	4.505
Macro Average	0,86	0,90	0,87	4.505
Weighted Average	0,90	0,88	0,88	4.505

Sumber: Diolah oleh penulis (2026)

Berdasarkan Tabel 11, hasil evaluasi model Multinomial Naïve Bayes pada *imbalanced dataset* menunjukkan akurasi sebesar 88,15% dari total 4.505 data testing. Pada kelas negatif, model memperoleh *precision* sebesar 0,76, *recall* sebesar 0,94, dan *F1-score* sebesar 0,84 dengan *support* sebanyak 1.510 data. Sementara itu, pada kelas positif diperoleh *precision* sebesar 0,96, *recall* sebesar 0,85, dan *F1-score* sebesar 0,91 dengan support sebanyak 2.995 data. Nilai macro average sebesar 0,86 untuk *precision*, 0,90 untuk *recall*, dan 0,87 untuk *F1-score* menunjukkan adanya perbedaan performa antar kelas, sedangkan *weighted average* sebesar 0,90, 0,88, dan 0,88 mencerminkan kinerja model secara keseluruhan dengan mempertimbangkan proporsi masing-masing kelas. Hasil tersebut menunjukkan bahwa model memiliki kemampuan yang baik dalam mengenali ulasan negatif, tetapi ketidakseimbangan distribusi kelas masih memengaruhi konsistensi performa model antar kelas.

VII. SIMPULAN

Hasil penelitian analisis sentimen terhadap ulasan pengguna ShopeePay di Google Play Store menggunakan algoritma Multinomial Naïve Bayes menunjukkan bahwa distribusi kelas berpengaruh terhadap kinerja model klasifikasi. Penelitian ini menggunakan dua skenario dengan jumlah data yang sama, yaitu 30.032 ulasan. Skenario pertama menggunakan balanced dataset yang terdiri atas 15.016 ulasan positif dan 15.016 ulasan negatif, sedangkan skenario kedua menggunakan imbalanced dataset yang terdiri atas 20.016 ulasan positif dan 10.016 ulasan negatif.

Pada balanced dataset, model Multinomial Naïve Bayes menghasilkan accuracy sebesar 91,03%. Kinerja model pada kedua kelas juga relatif seimbang. Kelas negatif memperoleh precision 0,90, recall 0,92, dan F1-score 0,91, sedangkan kelas positif memperoleh precision 0,92, recall 0,90, dan F1-score 0,91. Hasil tersebut menunjukkan bahwa model mampu membedakan sentimen positif dan negatif secara konsisten ketika distribusi jumlah data pada kedua kelas berada dalam kondisi seimbang.

Pada imbalanced dataset yang ditangani menggunakan SMOTE, model menghasilkan accuracy sebesar 88,15%. Kelas negatif memperoleh precision 0,76, recall 0,94, dan F1-score 0,84, sedangkan kelas positif memperoleh precision 0,96, recall 0,85, dan F1-score 0,91. Tingginya recall pada kelas negatif menunjukkan bahwa model mampu mengenali sebagian besar data pada kelas minoritas setelah penerapan SMOTE. Namun, rendahnya precision kelas negatif menunjukkan bahwa masih terdapat cukup banyak ulasan positif yang diprediksi sebagai negatif. Dengan demikian, SMOTE membantu meningkatkan representasi kelas minoritas, tetapi belum mampu menghasilkan keseimbangan performa antar kelas seperti pada balanced dataset.

Dalam konteks penelitian ini, keterbatasan SMOTE terlihat dari kemampuannya yang hanya menambah sampel sintetis berdasarkan pola data minoritas yang telah tersedia. Sampel sintetis tersebut tidak selalu menghasilkan informasi baru yang mampu memperkuat pemisahan karakteristik antara kelas positif dan negatif. Akibatnya, meskipun jumlah data training menjadi lebih seimbang, performa model belum tentu meningkat secara keseluruhan. Hal ini terlihat dari accuracy sebesar 88,15% yang masih lebih rendah dibandingkan balanced dataset sebesar 91,03%, serta adanya ketimpangan precision dan recall pada kelas negatif. Oleh karena itu, penerapan SMOTE tidak dapat

secara otomatis menjamin peningkatan akurasi model, khususnya ketika karakteristik data sintetis belum mampu merepresentasikan variasi bahasa dan pola sentimen kelas minoritas secara optimal.

Hasil analisis teks juga menunjukkan bahwa ulasan positif banyak memuat kata seperti bagus, mudah, transaksi, promo, cepat, dan ShopeePay, yang mencerminkan apresiasi pengguna terhadap kemudahan penggunaan, kecepatan transaksi, dan program promosi. Sebaliknya, ulasan negatif banyak berkaitan dengan kata seperti saldo, login, akun, payment, error, verifikasi, dan top-up, yang mengindikasikan adanya keluhan terkait transaksi, keterlambatan saldo, proses verifikasi akun, serta kendala teknis aplikasi.

Secara keseluruhan, Multinomial Naïve Bayes mampu digunakan secara efektif untuk mengklasifikasikan sentimen ulasan pengguna ShopeePay, dengan performa terbaik diperoleh pada kondisi data yang seimbang. Penerapan SMOTE mampu membantu mengurangi ketimpangan kelas dan meningkatkan representasi kelas minoritas, tetapi belum menghasilkan performa yang lebih baik dibandingkan dataset yang telah seimbang sejak awal. Temuan ini menegaskan bahwa penyeimbangan kelas tidak hanya bergantung pada jumlah data, tetapi juga pada kualitas dan keterwakilan karakteristik data pada masing-masing kelas. Bagi ShopeePay, hasil penelitian dapat menjadi dasar untuk memprioritaskan perbaikan pada keandalan transaksi, pengelolaan saldo, proses verifikasi akun, serta kendala teknis aplikasi.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Universitas Muhammadiyah Sidoarjo, khususnya Program Studi Bisnis Digital, atas dukungan akademik dan fasilitas yang diberikan selama proses pelaksanaan penelitian ini. Penulis juga menyampaikan apresiasi kepada dosen pembimbing serta seluruh pihak yang telah memberikan arahan, masukan, dan dukungan dalam proses pengumpulan, pengolahan, analisis data, hingga penyusunan artikel ini. Semoga hasil penelitian mengenai analisis sentimen ulasan pengguna ShopeePay ini dapat memberikan manfaat bagi pengembangan penelitian di bidang bisnis digital, *machine learning*, dan analisis sentimen serta menjadi referensi bagi penelitian selanjutnya.

Conflict of Interest Statement:
The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.