

Analisis Sentimen Ulasan Shopee di Google Play Store Menggunakan Metode Multinomial Naïve Bayes

Devina Widi Echa, Prastyadi Wibawa Rahayu*, I Made Dwi Ardiada

Informatika, Fakultas Teknologi dan Informatika, Universitas Dhyana Pura, Bali, Indonesia

Email: ¹22121101019@undhirabali.ac.id, ^{2*}prastyadiwibawa@undhirabali.ac.id, ³dwiardiada@undhirabali.ac.id

Submitted: 08/06/2026; Accepted: 02/07/2026; Published: 30/09/2026

Abstrak— Pertumbuhan pesat di sektor perdagangan elektronik telah menempatkan belanja daring sebagai pilar ekonomi digital yang krusial. Namun, tingginya interaksi pengguna pada platform tersebut menghasilkan jutaan ulasan yang mengandung opini subjektif mengenai kualitas layanan serta produk. Besarnya volume data tersebut menyulitkan pihak pengelola untuk melakukan identifikasi kepuasan pelanggan secara manual dan cepat. Permasalahan utama dalam penelitian ini adalah bagaimana mengotomatisasi proses klasifikasi sentimen pengguna agar pengelola dapat memahami umpan balik pelanggan secara efisien. Tujuan penelitian ini adalah melakukan analisis sentimen ulasan pengguna pada platform belanja daring menggunakan metode *Multinomial Naïve Bayes Classifier* sebagai solusi klasifikasi otomatis. Peneliti menggunakan data latih 80% dari total dataset sebanyak 44.481 dan data uji mencakup 20% dari total dataset sebanyak 11.121 jika ditambahkan kedua data dihasilkan 55.602 data cleaning. Hasil penelitian menunjukkan tingkat akurasi sebesar 87.22%, *precision* sebesar 85.21%, *recall* sebesar 87.22% dan *f1-score* sebesar 85.76%. Pembahasan menunjukkan algoritma memiliki performa tinggi dalam mengenali opini positif dan negatif, meskipun memerlukan optimasi untuk kelas netral. Secara praktis, penelitian ini membantu pengembang dalam memperbaiki sistem layanan secara berkelanjutan. Penelitian selanjutnya disarankan menerapkan teknik penyeimbangan data guna meningkatkan sensitivitas model.

Kata Kunci: Analisis Sentimen; Perdagangan Elektronik; Pembelajaran Mesin; *Multinomial Naïve Bayes*; Ulasan Pengguna; Klasifikasi Teks

Abstract— The rapid growth of the e-commerce sector has positioned online shopping as a crucial pillar of the digital economy. However, the high level of user interaction on these platforms generates millions of reviews containing subjective opinions regarding service and product quality. The large volume of review data makes it difficult for platform providers to identify customer satisfaction manually and efficiently. The main problem addressed in this study is how to automate the sentiment classification process so that platform providers can understand customer feedback more effectively. Therefore, this study aims to perform sentiment analysis of user reviews on an online shopping platform using the Multinomial Naïve Bayes Classifier as an automated classification method. The dataset consisted of 55,602 cleaned reviews, which were divided into 44,481 training data (80%) and 11,121 testing data (20%). The results showed an accuracy of 87.22%, a precision of 85.21%, a recall of 87.22%, and an F1-score of 85.76%. The findings indicate that the proposed algorithm achieved high performance in classifying positive and negative sentiments but showed lower performance in identifying neutral sentiments due to class imbalance in the dataset. Practically, this study provides useful insights for developers to continuously improve service quality based on customer feedback. Future research is recommended to apply data balancing techniques to improve the model's performance, particularly in classifying the minority class.

Keywords: Sentiment Analysis; E-commerce; Machine Learning; Multinomial Naïve Bayes; User Reviews; Text Classification

1. PENDAHULUAN

Pertumbuhan ekonomi yang pesat telah menempatkan *e-commerce* sebagai pilar ekonomi digital yang krusial di Indonesia. Persaingan bisnis yang ketat antar platform menuntut setiap perusahaan untuk terus berinovasi. *Shopee*, sebagai salah satu pemimpin pasar, terus menghadapi tantangan seiring dengan meningkatnya interaksi pengguna dari berbagai kalangan [1]. Meskipun *Shopee* berperan penting dalam dinamika pasar, kompleksitas sistemnya yang mencakup berbagai fitur, iklan, dan keterlibatan logistic sering kali menciptakan hambatan dalam pengalaman pengguna. Dalam ekosistem ini, pembeli memiliki kebebasan untuk membandingkan harga, namun mereka juga rentan terhadap risiko ketidaksesuaian produk atau layanan. Untuk memitigasi risiko tersebut, pengguna sering memberikan peringkat (*rating*) sebagai umpan balik, di mana *rating* 1-2 mencerminkan ketidakpuasan, 3 mencerminkan netral, dan 4-5 menunjukkan kepuasan [2]. Tantangan utama yang dihadapi adalah volume data ulasan yang mencapai puluhan juta, yang jika tidak dikelola dengan metode terstruktur, hanya akan menjadi data mentah yang tidak memberikan nilai informatif bagi pengelola. Oleh karena itu, penelitian ini menawarkan solusi penerapan analisis sentimen berbasis *text mining*, yang mencakup tahap *text processing* dan pemodelan klasifikasi menggunakan *Multinomial Naïve Bayes*. Metode ini dipilih karena efisiensinya dalam menangani klasifikasi teks dengan asumsi fitur yang independen, yang terbukti mampu memberikan tingkat presisi yang baik dalam mengklasifikasikan komentar positif, negatif, dan netral [3].

Penelitian analisis sentimen telah berkembang pesat, namun tinjauan kritis terhadap literatur terkini mengungkapkan adanya inkonsistensi metodologis yang serius. Meskipun studi oleh Zulcharnain et al. (2025) [4] menegaskan efektivitas *Multinomial Naïve Bayes*, terdapat variasi performa yang fluktuatif di berbagai studi terdahulu, mulai dari akurasi 71% pada Aplikasi yang sama yakni *shopee* [5]. Hingga akurasi 88% [6]. Inkonsistensi ini tidak hanya sekadar mencerminkan perbedaan algoritma, melainkan menunjukkan kelemahan mendasar dalam standar *preprocessing* dan kurangnya validasi terhadap volume data yang heterogen. Mayoritas studi yang membandingkan metode, seperti penggunaan *Random Forest* atau kombinasi *TF-IDF* dan *LSTM*, sering kali terjebak pada pengabaian distribusi data yang sebenarnya [6], [7]. Celah riset (*research gap*) yang krusial terletak pada kurangnya atensi terhadap *class imbalance*. Banyak penelitian mengklaim akurasi tinggi tanpa menyadari bahwa nilai tersebut didominasi oleh kelas mayoritas (positif/negatif), sementara kelas netral sering kali diabaikan atau justru terklasifikasi secara salah akibat ketidakmampuan model dalam menangkap ambiguitas linguistic [8]. Penelitian tahun 2026 yang menggunakan *Naïve Bayes* dengan akurasi 84,88% pun belum sepenuhnya menyelesaikan problem sensitivitas model pada kategori netral. Dengan demikian, tantangan utama saat ini bukanlah pada pencarian algoritma baru, melainkan pada optimasi strategi penanganan ketidakseimbangan kelas dan penguatan fitur pada teks tidak terstruktur berskala besar, yang hingga kini masih menjadi titik buta dalam narasi efektivitas analisis sentimen.

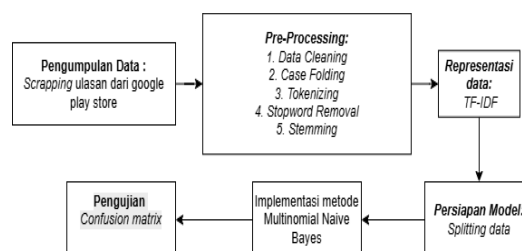
Tujuan penelitian ini adalah untuk melakukan analisis sentimen terhadap ulasan pengguna aplikasi Shopee di *Google Play Store* menggunakan *Multinomial Naïve Bayes Classifier*, serta mengukur kinerja model secara komprehensif melalui *confusion matrix*. Harapan yang ingin dicapai adalah terbentuknya model klasifikasi yang tangguh dalam mendeteksi persepsi pengguna secara akurat. Hasil penelitian ini diharapkan dapat menjadi acuan bagi pihak terkait dalam memahami keluhan pengguna untuk perbaikan sistem, sekaligus memperkaya literatur akademik mengenai penerapan *text mining* pada data *e-commerce* berskala besar.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Pada penelitian ini, penulis melakukan analisis sentimen Aplikasi *shopee* pada *google play store* menggunakan metode *multinomial naïve bayes*. Sumber data yang digunakan merupakan ulasan pengguna Aplikasi *shopee* di *google play store*. Data penelitian diperoleh dari ulasan pengguna aplikasi *Shopee* di *Google Play Store*. Pengumpulan data dilakukan menggunakan teknik *web scraping* dengan bantuan *library Python* yaitu *google-play-scraper*. Proses *scraping* dilaksanakan pada tanggal 20 April 2026 dan menghasilkan 68.301 data ulasan mentah yang mencakup periode ulasan dari 1 Desember 2025 hingga 30 Januari 2026. Kriteria ulasan yang diambil adalah ulasan dalam Bahasa Indonesia untuk menjaga konsistensi konteks sentimen. Untuk menjaga representasi proporsi kelas yang seimbang dalam set pelatihan dan pengujian, penelitian ini menggunakan *Stratified Split* dengan rasio 80:20. Penggunaan metode *stratified* memastikan bahwa distribusi kelas (positif, netral, negatif) pada data latih dan data uji memiliki proporsi yang sama dengan dataset asli, sehingga meminimalisir bias [9]. Sentimen yang digunakan penulis merupakan sentimen negatif dengan *rating* 1-3, netral dengan *rating* 3, dan positif dengan *rating* 4-5 [10].

Sebelum dilakukan klasifikasi, teks melalui tahap pembersihan (*text cleaning*), tokenisasi, penghapusan kata umum (*stopword removal*), dan *stemming*. Selanjutnya, teks dikonversi ke dalam representasi numerik menggunakan *TF-IDF* (*Term Frequency-Inverse Document Frequency*) dengan konfigurasi *N-gram Unigram* (1,1) untuk menangkap fitur kata tunggal, *Max Features* dibatasi pada 5.000 fitur teratas untuk mengurangi dimensi dan *noise*, dan Pembobotan Menggunakan skema *sublinear tf scaling* untuk menekan pengaruh frekuensi kata yang terlalu tinggi. Tujuan penelitian yakni untuk mengetahui hasil tingkat akurasi metode *Multinomial Naïve Bayes Classifier* pada ulasan *shopee* di *google play store* dalam konteks komentar Bahasa Indonesia. Peneliti melakukan seluruh proses menggunakan *tools google collabs*, tahapan yang dilakukan penulis yakni *scrapping* ulasan, *data cleaning*, *case folding*, *tokenizing*, *stopword removal*, *stemming*, *splitting data*, implementasi *multinomial naïve bayes*, *confusion matrix*. Seluruh proses Disusun sistematis pada gambar 1 untuk memastikan penelitian terstruktur.



Gambar 1. Rancangan Penelitian

2.2 Scrapping data

Tahap *Scrapping* merupakan tahap dalam penelitian ini berfokus pada data media sosial, yang dianggap sebagai kanal paling dinamis dalam merefleksikan perilaku manusia serta fenomena terkini [10].

Tabel 1. Ringkasan *Dataset*

Keterangan Data	Jumlah
Total Komentar Awal	68.301
Tahap Cleaning Data	55.602

2.3 Pre-Processing

Tahap *Pre-Processing* peneliti melakukan 5 tahap proses yakni *data cleaning*, *case folding*, *tokenizing*, *stopword removal*, *stemming*. *Pre-Processing* adalah tahapan pra-pemrosesan bertujuan mentransformasikan data teks yang tidak terstruktur menjadi format yang lebih bersih dan terorganisir [11]. Berikut tahapan yang dilakukan :

- Data Cleaning*, untuk membersihkan ulasan dari simbol, angka, *URL*, hashtag [12].
- Case Folding*, untuk mengubah seluruh huruf menjadi huruf kecil (*lowercase*) [12].
- Tokenizing*, berfungsi memecah kalimat menjadi potongan kata-kata[13].
- Stopword Removal*, menyeleksi kata yang relevan guna memperjelas sentiment terhadap objek [14].
- Stemming*, proses memperoleh kata dasar [15].

2.4 TF-IDF

TF-IDF (*Term Frequency - Inverse Document Frequency*) merupakan metode yang digunakan untuk pembobotan kata. Metode ini memberikan bobot pada setiap kata dalam deskripsi, dimana istilah yang unik dan informatif akan diprioritaskan dengan nilai yang lebih tinggi [16].

$$TF-IDF = TF * IDF \quad (1)$$

2.5 Splitting Data

Splitting data merupakan proses pembagian data ke dalam kelompok data latih dan data uji. Tujuan utamanya adalah Langkah ini bertujuan untuk mengevaluasi performa model menggunakan data uji yang tidak terlibat dalam proses pelatihan. Peneliti menggunakan data latih 80% dari total dataset sebanyak 44.481 dan data uji mencakup 20% dari total dataset sebanyak 11.121 jika ditambahkan kedua data dihasilkan 55.602 data *cleaning* [17].

2.6 Implementasi Multinomial Naïve Bayes

Naïve Bayes merupakan metode dalam pembelajaran mesin yang sering digunakan untuk klasifikasi teks. Metode ini menganggap bahwa setiap fitur bersifat independen satu sama lain secara kondisional. Dengan menghitung peluang (*probabilitas*) suatu data termasuk dalam kelas tertentu berdasarkan fitur-fitur yang digunakan. Disebut sebagai “*Naïve*” karena mengasumsikan setiap fitur bersifat independen satu sama lain, dalam kenyataan tidak selalu demikian [3].

$$P(H|X) = \frac{P(X|H).P(H)}{P(X)} \quad (2)$$

2.7 Confusion matrix

Confusion matrix merupakan metode evaluasi model klasifikasi yang berbentuk tabel *matrix* untuk membandingkan hasil prediksi model dengan data aktual. Pada matriks terdapat gambar detail mengenai prediksi benar atau salah untuk setiap kelasnya [18].

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

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

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

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

3. HASIL DAN PEMBAHASAN

3.1 Pengumpulan data (*Scrapping Data*)

Pada tahap ini, peneliti melakukan proses *scraping* data ulasan aplikasi Shopee melalui situs *com.shopee.id* pada tanggal 20 April 2026, yang berhasil mengumpulkan total 68.301 entri data ulasan dalam Bahasa Indonesia. Data yang diperoleh mencakup berbagai atribut mendalam, seperti *reviewId* sebagai identitas unik ulasan, *userName*,

userImage, isi komentar (*content*), serta *score* rating yang diberikan oleh pengguna. Selain itu, dataset ini memuat metrik interaksi seperti *thumbsUpCount* untuk jumlah dukungan, dimensi waktu melalui *at* dan *repliedAt* untuk mencatat durasi respons, serta informasi teknis berupa *replyContent*, *appVersion*, dan *reviewCreated* yang memberikan gambaran komprehensif mengenai pengalaman pengguna serta efektivitas respons pihak pengembang terhadap kendala yang dilaporkan, sebagaimana disajikan secara terperinci dalam struktur data berikut.

3.2 Pre-Processing

3.2.1 Data Cleaning

Data yang telah dilakukan pembersihan data (*cleaning data*) dari ulasan yang dilakukan *scrapping* 68.301 ulasan, setelah dilakukan *cleaning* menjadi 55.602 ulasan. Pada tahap ini dilakukan *handling missing value* untuk menangani data yang kosong atau hilang. Langkah krusial ini dalam tahap *pra-pemrosesan* untuk memastikan integritas dan validitas *dataset* sebelum diolah ke tahap analisis sentimen. Melalui perbaikan struktur data yang tidak konsisten dan penyesuaian kembali antara identitas baris dengan isi ulasan serta skor rating, metode ini berhasil meminimalisir potensi *bias* yang diakibatkan oleh *noise* atau pergeseran data pada *dataset* mentah. Implementasi pembersihan ini menjadi sangat vital bagi akurasi model penelitian, mengingat keakuratan hasil analisis bergantung pada kualitas data yang terstruktur secara sistematis dan bersih dari kesalahan format.

3.2.2 Case Folding

Proses *case folding* dilakukan agar kata yang kapital menjadi huruf kecil misalnya “Shopee”, “SHOPEE” akan dilakukan *case folding* Sehingga menjadi kata yang memiliki huruf kecil yakni “shopee”. Setelah dilakukan *running* dari *code*, *output* berupa data CSV berupa kalimat yang sebelum dan sesudah dilakukan *case folding* pada gambar dibawah ini :

Tabel 2. Hasil Proses Case Folding

No	content	Label	text_clean
1	sangat membantu sekali	Positif	sangat membantu sekali
2	rekomendasinya GK sesuai kebutuhan	Negatif	rekomendasinya gk sesuai kebutuhan
3	shopee Lo bisah diem ngakk gw lagi enak enak baca Wattpad trus tiba tiba ada iklan Lo ganggu tau ngakk sistemnya tolong di perbaiki	Negatif	shopee lo bisah diem ngakk gw lagi enak enak baca wattpad trus tiba tiba ada iklan lo ganggu tau ngakk sistemnya tolong di perbaiki
4	tolong dong shopee iklannya dikondisikan nyebelin banget ga pencet apapun tiba-tiba dialihkan ke shopee mulu lagi enak baca wattpad nih tiba-tiba ke shopee padahal ga pencet apapun bikin sebel tau nggak sih jadinya saya tau tujuan nya untuk promosi tapi ya jangan mengganggu kenyamanan ketika buka apk lain dong tolong untuk masalah periklanan segera diatasi bikin bete soalnya	Netral	tolong dong shopee iklannya dikondisikan nyebelin banget ga pencet apapun tibatiba dialihkan ke shopee mulu lagi enak baca wattpad nih tibatiba ke shopee padahal ga pencet apapun bikin sebel tau nggak sih jadinya saya tau tujuan nya untuk promosi tapi ya jangan mengganggu kenyamanan ketika buka apk lain dong tolong untuk masalah periklanan segera diatasi bikin bete soalnya
5	oke sip	Positif	oke sip
6	selama pake shoppy ad beberapa si yank hasil barang nya mengecewakan tapi sebagian besar itu barang nya bagus semua	Positif	selama pake shoppy ad beberapa si yank hasil barang nya mengecewakan tapi sebagian besar itu barang nya bagus semua
7	paraaah aplikasinya	Negatif	paraaah aplikasinya
8	sangat sangat membantu	Positif	sangat sangat membantu
9	barang2 TDK mengecewakan.. dan kulir2 baik	Positif	barang2 tdk mengecewakan dan kulir 2 baik
10	man tap	Positif	man tap

3.2.3 Tokenizing

Proses *tokenizing* dilakukan untuk mengubah data teks yang telah dibersihkan menjadi daftar kata (*list of tokens*) yang siap dikonversi ke bentuk numerik. Meskipun tahapan ini telah berhasil menstandarisasi *input* teks, perlu dicatat secara kritis bahwa persistensi rendahnya performa pada kelas netral seperti yang teridentifikasi dalam pengujian kemungkinan besar berakar dari kurangnya representasi fitur netral yang *distingtif* pada tahap ini.

Dengan jumlah sampel netral yang sangat sedikit. Berikut hasil stopword yang dilanjutkan dengan proses tokenizing dibawah ini:

Tabel 3. Hasil Proses *Tokenizing* sebelum dan sesudah

No	content	text_tokens
1	sangat membantu sekali	['membantu']
2	rekomendasinya GK sesuai kebutuhan	['rekomendasinya', 'gk', 'sesuai', 'kebutuhan']
3	shopeeee Lo bisah diem ngakk gw lagi enak enak baca Wattpad trus tiba tiba ada iklan Lo ganggu tau ngakk sistemnya tolong di perbaiki	['shopee', 'lo', 'bisah', 'diem', 'ngakk', 'gw', 'enak', 'enak', 'baca', 'wattpad', 'trus', 'iklan', 'lo', 'ganggu', 'tau', 'ngakk', 'sistemnya', 'tolong', 'perbaiki']
4	tolong dong shopee iklannya dikondisikan nyebelin banget ga pencet apapun tiba-tiba dialihkan ke shopee mulu lagi enak baca wattpad nih tiba-tiba ke shopee padahal ga pencet apapun bikin sebel tau nggak sih jadinya saya tau tujuan nya untuk promosi tapi ya jangan mengganggu kenyamanan ketika buka apk lain dong tolong untuk masalah periklanan segera diatasi bikin bete soalnya	['tolong', 'shopee', 'iklannya', 'dikondisikan', 'nyebelin', 'banget', 'ga', 'pencet', 'apapun', 'tiba-tiba', 'dialihkan', 'shopee', 'mulu', 'enak', 'baca', 'wattpad', 'nih', 'tiba-tiba', 'shopee', 'ga', 'pencet', 'apapun', 'bikin', 'sebel', 'tau', 'nggak', 'sih', 'tau', 'tujuan', 'nya', 'promosi', 'ya', 'mengganggu', 'kenyamanan', 'buka', 'apk', 'tolong', 'periklanan', 'diatasi', 'bikin', 'bete']
5	oke sip	['oke', 'sip']
6	selama pake shoppy ad beberapa si yank hasil barang nya mengecewkan tapi sebagian besar itu barang nya bagus semua	['pake', 'shoppy', 'ad', 'si', 'yank', 'hasil', 'barang', 'nya', 'mengecewkan', 'barang', 'nya', 'bagus']
7	paraaah aplikasinya	['paraaah', 'aplikasinya']
8	sangat sangat membantu	['membantu']
9	barang2 TDK menceewKan.. dan kulir2 baik	['barang', 'tdk', 'menceewkan', 'kulir']
10	man tap	['man', 'tap']
11	sangat memudahkan setiap transaksi	['memudahkan', 'transaksi']

3.2.4 Stopword Removal

Tahap selanjutnya merupakan *stopword removal* menggunakan *library NLTK*, proses eliminasi kata-kata umum yang sering muncul dan tidak memiliki pengaruh signifikan, seperti “yang”, “dan”, “di”, “dari” akan dihapus otomatis oleh algoritma. Namun, secara kritis, perlu diakui bahwa proses penyaringan ini memiliki dampak ganda. Di satu sisi, ia membersihkan data dari *noise*, namun di sisi lain, ia berisiko menghilangkan kata-kata penghubung yang sebenarnya mungkin mengandung konteks penting bagi sentimen yang bersifat ambigu. Berikut Sebelum dilakukan *stopword* dan hasilnya setelah dilakukan *text stopwords* :

Tabel 4. Hasil Proses Stopword Removal

No.	content	text_StopWord
1	sangat membantu sekali	membantu
2	rekomendasinya GK sesuai kebutuhan	rekomendasinya gk sesuai kebutuhan
3	shopeeeee Lo bisah diem ngakk gw lagi enak enak baca Wattpad trus tiba tiba ada iklan Lo ganggu tau ngakk sistemnya tolong di perbaiki	shopee lo bisah diem ngakk gw enak enak baca wattpad trus iklan lo ganggu tau ngakk sistemnya tolong perbaiki
4	tolong dong shopee iklannya dikondisikan nyebelin banget ga pencet apapun tiba-tiba dialihkan ke shopee mulu lagi enak baca wattpad nih tiba-tiba ke shopee padahal ga pencet apapun bikin sebel tau nggak sih jadinya saya tau tujuan nya untuk promosi tapi ya jangan mengganggu kenyamanan ketika buka apk lain dong tolong untuk masalah periklanan segera diatasi	tolong shopee iklannya dikondisikan nyebelin banget ga pencet apapun tibatiba dialihkan shopee mulu enak baca wattpad nih tibatiba shopee ga pencet apapun bikin sebel tau nggak sih taujuan nya promosi ya mengganggu kenyamanan buka apk tolong periklanan diatasi bikin bete

5	oke sip	oke sip
6	selama pake shoppy ad beberapa si yank hasil barang nya mengecewakan tapi sebagian besar itu barang nya bagus semua	pake shoppy ad si yank hasil barang nya mengecewakan barang nya bagus
7	paraaah aplikasinya	paraaah aplikasinya
8	sangat sangat membantu	membantu
9	barang2 TDK mencewkan.. dan kulir2 baik	barang tdk mencewkan kulir
10	man tap	man tap
11	sangat memudahkan setiap transaksi	memudahkan transaksi

3.2.5 Stemming

Setelah proses *stopword removal* selesai, tahap selanjutnya adalah *stemming* untuk mereduksi kata ke dalam bentuk dasarnya. Proses ini krusial dalam menyamakan variasi imbuhan kata agar model dapat mengenali akar kata yang sama sebagai satu fitur yang konsisten. Pada penelitian ini menggunakan *library* *sastawi*. Berikut hasil ulasan sebelum dilakukan stemming dan setelah di ubah ke dalam *stemming* :

Tabel 5. Hasil Proses Stemming

No.	content	text_steamindo
1	sangat membantu sekali	bantu
2	rekomendasinya GK sesuai kebutuhan	rekomendasi gk sesuai butuh
3	shopeeee Lo bisah diem ngakk gw lagi enak enak baca Wattpad trus tiba tiba ada iklan Lo ganggu tau ngakk sistemnya tolong di perbaiki	shopee lo bisah diem ngakk gw enak enak baca wattpad trus iklan lo ganggu tau ngakk sistem tolong baik
4	tolong dong shopee iklannya dikondisikan nyebelin banget ga pencet apapun tiba-tiba dialihkan ke shopee mulu lagi enak baca wattpad nih tiba-tiba ke shopee padahal ga pencet apapun bikin sebel tau nggak sih jadinya saya tau tujuan nya untuk promosi tapi ya jangan mengganggu kenyamanan ketika buka apk lain dong tolong untuk masalah periklanan segera diatasi bikin bete soalnya	tolong shopee iklan kondisi nyebelin banget ga pencet apa tibatiba alih shopee mulu enak baca wattpad nih tibatiba shopee ga pencet apa bikin sebel tau nggak sih tauju nya promosi ya ganggu nyaman buka apk tolong iklan atas bikin bete
5	oke sip	oke sip
6	selama pake shoppy ad beberapa si yank hasil barang nya mengecewakan tapi sebagian besar itu barang nya bagus semua	pake shoppy ad si yank hasil barang nya mengecewakan barang nya bagus
7	paraaah aplikasinya	paraah aplikasi
8	barang2 TDK mencewkan.. dan kulir2 baik	barang tdk mencewkan kulir
9	man tap	man tap
10	sangat memudahkan setiap transaksi	mudah transaksi
11	mau cara apa aja pasti ada dan ongkirnya murah banget	aja ongkirnya murah banget

3.3 TF-IDF

Pada tahap ini, metode TF-IDF digunakan untuk mengubah data tekstual menjadi representasi numerik sebelum diproses oleh algoritma klasifikasi. Proses tersebut dilakukan menggunakan fungsi *fit_transform()* pada data latih dan *transform()* pada data uji sehingga dimensi fitur yang dihasilkan tetap konsisten. Hasil pengecekan menggunakan atribut *.shape* menunjukkan bahwa data pelatihan terdiri atas 44.481 ulasan, sedangkan data pengujian berjumlah 11.121 ulasan. Secara analitis, penggunaan TF-IDF memiliki keterbatasan dalam penelitian ini karena tidak mampu mengatasi ketidakseimbangan distribusi kelas pada dataset. Distribusi ulasan yang didominasi oleh kelas positif dan negatif menyebabkan bobot TF-IDF lebih banyak merepresentasikan karakteristik kedua kelas tersebut, sedangkan kelas netral yang memiliki jumlah data sangat sedikit tidak memperoleh representasi fitur yang memadai. Akibatnya, fitur-fitur yang dihasilkan untuk kelas netral memiliki daya pembeda (*discriminative power*) yang rendah sehingga model sering mengalami kesalahan dalam mengklasifikasikan ulasan netral ke dalam kelas mayoritas. Kondisi ini sejalan dengan hasil evaluasi model yang menunjukkan performa kelas netral jauh lebih rendah dibandingkan kelas positif dan negatif, dengan jumlah prediksi benar yang sangat terbatas. Oleh karena itu, rendahnya performa pada kelas netral merupakan salah satu keterbatasan penelitian ini yang terutama dipengaruhi oleh ketidakseimbangan data, bukan semata-mata oleh algoritma klasifikasi yang digunakan. Permasalahan tersebut dapat diminimalkan melalui penerapan teknik

penyeimbangan data, seperti *oversampling* atau *undersampling*, maupun penggunaan metode ekstraksi fitur yang lebih kontekstual agar representasi kelas minoritas menjadi lebih optimal.

3.4 Splitting Data

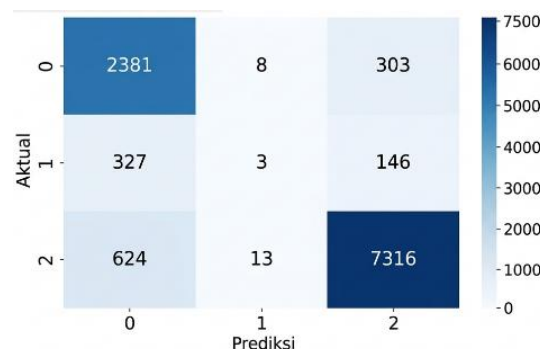
Setelah dilakukan *TF-IDF* dilanjutkan dengan data *splitting* menjadi data *training* dan data *testing* untuk mengevaluasi kinerja model *machine learning* secara objektif. Berikut *source code* yang digunakan untuk *splitting model*. Data yang digunakan pada penelitian ini terdiri dari data latih (*training data*) dan data uji (*testing data*) dengan rasio pembagian 80:20. Dari total dataset sebanyak 55.602 ulasan, sebanyak 44.481 ulasan (80%) digunakan sebagai data latih untuk melatih algoritma *Multinomial Naïve Bayes* dalam mengenali pola kata pada setiap kategori sentimen. Sementara itu, 11.121 ulasan sisanya (20%) digunakan sebagai data uji untuk mengevaluasi performa model.

Tabel 6. Distribusi Dataset

Keterangan Data	Jumlah
Data Uji (20%)	11.121
Data Latih (80%)	44.481
Total Data (100%)	55.602

3.5 Implementasi Multinomial Naïve Bayes

Setelah tahap *preprocessing* selesai, dilanjutkan dengan teks yang ditransformasi menjadi numerik menggunakan teknik *count vector* untuk merepresentasikan frekuensi kemunculan setiap kata dalam bentuk *vector* angka agar dapat diproses. Data dibagi menjadi data *training* dan data uji. Data *training* digunakan untuk model klasifikasi sentimen algoritma *Multinomial Naïve Bayes* melalui fungsi *nb.fit*. *Fiting* model menggunakan algoritma *Multinomial Naïve Bayes* pada representasi *count vector* ini memiliki kerentanan tinggi terhadap masalah ketidakseimbangan kelas (*class imbalance*). Karena algoritma ini bekerja berdasarkan perhitungan probabilitas kemunculan kata pada masing-masing kelas, jumlah data netral yang sangat sedikit menyebabkan variasi dan bobot frekuensi kata pada kelas tersebut menjadi tidak signifikan dibandingkan kelas mayoritas. Akibatnya, fungsi *nb.fit* akan menghasilkan nilai probabilitas *apriori* dan probabilitas bersyarat yang jauh lebih condong ke kelas positif atau negatif. Keterbatasan matematis dari model berbasis probabilitas frekuensi inilah yang menyebabkan ulasan netral sering kali "tenggelam" dan salah terprediksi ke dalam kelas yang memiliki jumlah sampel lebih dominan.



Gambar 4. Evaluasi Model Klasifikasi

3.6 Confusion matrix

Tahap terakhir merupakan tahap mengukur kinerja algoritma *Multinomial naïve bayes* menggunakan data uji guna memvalidasi kemampuan *system* dalam memprediksi sentiment secara akurat. Implementasi *Confusion matrix* dan *Classification report* digunakan untuk memetakan distribusi hasil klasifikasi serta mengidentifikasi kesalahan prediksi secara mendetail, sehingga kualitas dan reliabilitas model dalam melakukan analisis sentimen ulasan *Shopee* dapat dibuktikan secara objektif. Berikut hasil *classification report* pada gambar di bawah ini :

Tabel 7. Hasil Evaluasi Model *Multinomial Naïve Bayes*

Metrik	Nilai
Accuracy	87.22%
Precision	85.21%
Recall	87.22%
F1-Score	85.76%

Tabel 8. Confusion Matrix *Multinomial Naïve Bayes*

Aktual \ Prediksi	Prediksi Negatif	Prediksi Netral	Prediksi Positif
Aktual Negatif	2381	8	303
Aktual Netral	327	3	146
Aktual Positif	624	13	7316

Dari hasil diatas menunjukkan bahwasannya model berhasil mengklasifikasikan *accuracy* 87.22%, *Precision* 85.21%, *Recall* 87,22%, dan *f1 Score* 85,76%. model juga berhasil melakukan prediksi pada ulasan negatif sebanyak 2381, ulasan netral sebanyak 3, dan ulasan positif sebanyak 7316. Perlu dicatat bahwa ketimpangan jumlah data pada kelas netral, sebagaimana terlihat pada visualisasi *Confusion matrix* diatas, merupakan refleksi dari sifat dasar ulasan pengguna pada platform *e-commerce*. Pengguna aplikasi cenderung memberikan ulasan ketika mereka mengalami pengalaman yang sangat memuaskan (positif) atau sangat mengecewakan (negatif), sehingga sentimen yang bersifat ekspresif lebih dominan dibandingkan sentimen netral yang cenderung bersifat informatif atau datar. Minimnya jumlah data netral ini bukan sekadar kendala teknis, melainkan representasi dari pola perilaku konsumen yang kurang memberikan perhatian pada aspek-aspek yang tidak memicu reaksi emosional kuat. Akibatnya, model kehilangan kesempatan untuk mempelajari pola fitur yang unik bagi kelas netral, yang secara langsung berdampak pada rendahnya tingkat keberhasilan prediksi pada kategori tersebut.

4. KESIMPULAN

Penelitian ini telah berhasil mengimplementasikan metode *Multinomial Naïve Bayes* untuk menganalisis sentimen ulasan pengguna aplikasi Shopee di *Google Play Store* menggunakan dataset berskala besar sebanyak 55.602 ulasan. Berdasarkan pengujian *confusion matrix*, model menghasilkan performa klasifikasi dengan tingkat *accuracy* 87,22%, *precision* 85,21%, *recall* 87,22%, dan *f1-score* 85,76%. Hasil tersebut menunjukkan bahwa algoritma *Multinomial Naïve Bayes* efektif dalam memproses ulasan yang dinamis dan tidak terstruktur, khususnya dalam membedakan sentimen positif dan negatif secara jelas. Namun, model masih menunjukkan performa yang kurang optimal dalam mengklasifikasikan kelas netral. Secara keseluruhan, sentimen positif mendominasi hasil klasifikasi yang mengindikasikan tingkat kepuasan tinggi pengguna terhadap layanan Shopee, sementara ulasan negatif dan netral dapat menjadi referensi berharga bagi pengembang untuk mengevaluasi keluhan pengguna. Penelitian ini memiliki beberapa keterbatasan yang perlu diperhatikan. Pertama, keterbatasan utama terletak pada penggunaan satu algoritma klasifikasi saja, sehingga belum terdapat komparasi kinerja dengan metode *machine learning* lainnya. Kedua, batasan rentang waktu pengumpulan data selama dua bulan menyebabkan sistem belum mampu menangkap perubahan tren sentimen pengguna dalam jangka panjang. Oleh karena itu, penelitian selanjutnya diharapkan dapat melibatkan perbandingan dengan algoritma lain serta memperluas durasi pengumpulan data untuk menghasilkan analisis yang lebih komprehensif.

REFERENCES

- [1] A. Nurian, M. S. Ma'arif, I. N. Amalia, and C. Rozikin, "Analisis Sentimen Pengguna Aplikasi Shopee pada Situs Google Play Menggunakan Naive Bayes Classifier," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 12, no. 1, 2024, doi: 10.23960/jitet.v12i1.3631.
- [2] S. F. Kadir and A. Fairuzabadi, "Analisis Sentimen Ulasan Shopee di Google Play dengan TF-IDF dan Logistic Regression," *RIGGS: Journal of Artificial Intelligence and Digital Business*, vol. 4, no. 2, pp. 7940–57945, Jul. 2025, doi: 10.31004/riggs.v4i2.2850.
- [3] D. Prasetya, N. Rahaningsih, R. D. Dana, and C. L. Rohmat, "Analisis Sentimen Pengguna Aplikasi Mybluebird Dengan Algoritma Naïve Bayes Di Playstore," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 13, no. 1, 2025, doi: 10.23960/jitet.v13i1.5687.
- [4] R. Zulchamain, G. Abdurrahman, and D. Daryanto, "Analisis Sentimen Ulasan Duolingo dengan Metode Algoritma Multinomial Naïve Bayes," *Jurnal Informatika dan Teknologi Pendidikan*, vol. 5, no. 1, pp. 1–16, 2025, doi: 10.59395/jitp.v5i1.113.
- [5] J. Khatib Sulaiman, A. Shopee Salman Alfaris, and U. Amikom Yogyakarta, "Komparasi Metode KNN dan Naive Bayes terhadap Analisis Sentimen Pengguna Aplikasi Shopee," *The Indonesian Journal of Computer Science*, vol. 12, no. 5, pp. 2023–2766, Oct. 2023, doi: 10.33022/IJCS.V12I5.3304.
- [6] Muhammad Zainal Abidin, Mochammad Alfian Rosid, and Ade Eviyanti, "Analisis Sentimen Pengguna Aplikasi Shopee Pada Google Play Store Menggunakan Metode Random Forest," *STORAGE: Jurnal Ilmiah Teknik dan Ilmu Komputer*, vol. 4, no. 3, pp. 265–276, Aug. 2025, doi: 10.55123/storage.v4i3.6027.
- [7] M. Musfiroh, A. Tholib, and Z. Arifin, "Analisis Sentimen Terhadap Ulasan Aplikasi Shopee di Google Play Store Menggunakan Metode TF-IDF dan Long Short-Term Memory," *Journal of Electrical Engineering and Computer (JEECOM)*, vol. 6, no. 2, pp. 371–381, Oct. 2024, doi: 10.33650/jeeecom.v6i2.8713.

- [8] D. Prabowo Wiguna and S. E. Ginting, "Analisis Sentimen E-Commerce di Indonesia dengan Algoritma Naive Bayes (Studi Kasus Pada Platform Shopee, Tokopedia, Bukalapak, dan Lazada)," *AICOM: Artificial Intelligence and Computing*, vol. 1, no. 04, pp. 109–114, Mar. 2026, doi: <https://jurnal.mifandimandiri.com/index.php/aicom/article/view/995>
- [9] S. Khoerunnisa, D. F. Shiddieq, and D. Nurhayati, "Penerapan Algoritma Naive Bayes dengan Teknik TF-IDF dan Cross Validation untuk Analisis Sentimen Terhadap Starlink," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 5, no. 2, pp. 566–577, Mar. 2025, doi: 10.57152/malcom.v5i2.1852.
- [10] E. Yuniar, D. S. Utsalinah, and D. Wahyuningsih, "Implementasi Scrapping Data Untuk Sentiment Analysis Pengguna Dompot Digital dengan Menggunakan Algoritma Machine Learning," *Jurnal Janitra Informatika dan Sistem Informasi*, vol. 2, no. 1, pp. 35–42, 2022, doi: 10.25008/janitra.v2i1.145.
- [11] K. Rahayu, V. Fitria, D. Septhya, R. Rahmaddeni, and L. Efrizoni, "Klasifikasi Teks untuk Mendeteksi Depresi dan Kecemasan pada Pengguna Twitter Berbasis Machine Learning: Text Classification for Detecting Depression and Anxiety among Twitter Users based on Machine Learning", *MALCOM*, vol. 3, no. 2, pp. 108–114, Sep. 2023, doi: <https://www.journal.irpi.or.id/index.php/malcom/article/view/780>
- [12] S. Shevira, I. M. A. D. Suarjaya, and P. W. Buana, "Pengaruh Kombinasi dan Urutan Pre-Processing pada Tweets Bahasa Indonesia," *JITTER: Jurnal Ilmiah Teknologi dan Komputer*, vol. 3, no. 2, p. 1074, 2022, doi: 10.24843/jtrti.2022.v03.i02.p06.
- [13] R. A. Fitrianto and A. S. Editya, "Klasifikasi Tweet Sarkasme Pada Platform X Menggunakan Bidirectional Encoder Representations from Transformers," *Jurnal Teknologi Dan Sistem Informasi Bisnis*, vol. 6, no. 3, pp. 366–371, 2024, doi: 10.47233/jteksis.v6i3.1344.
- [14] Diky, R. Astuti, and D. D. Saputra, "Optimasi Feature Selection Text Mining: Stemming dan Stopword untuk Sentimen Analisis Aplikasi SatuSehat," *Innovative: Journal Of Social Science Research*, vol. 4, no. 1, pp. 7537–7548, 2024, doi: <https://10.31004/innovative.v4i1.8759>.
- [15] A. Sinaga and S. P. Nainggolan, "Analisis Perbandingan Akurasi Dan Waktu Proses Algoritma Stemming Arifin-Setiono Dan Nazief-Adriani Pada Dokumen Teks Bahasa Indonesia," *Sebatik*, vol. 27, no. 1, pp. 63–69, 2023, doi: 10.46984/sebatik.v27i1.2072.
- [16] J. V. Santoti, J. Jocelyn, and H. Irsyad, "Implementasi Term Frequency - Inverse Document Frequency dan Cosine Similarity untuk Analisis Kemiripan Deskripsi Produk Halal," *Jurnal Software Engineering and Computational Intelligence*, vol. 3, no. 01, pp. 44–52, 2025, doi: 10.36982/jseci.v3i01.5421.
- [17] F. Putra, H. F. Tahiyat, R. M. Ihsan, R. Rahmaddeni, and L. Efrizoni, "Penerapan Algoritma K-Nearest Neighbor Menggunakan Wrapper Sebagai Preprocessing untuk Penentuan Keterangan Berat Badan Manusia," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 4, no. 1, pp. 273–281, 2024, doi: 10.57152/malcom.v4i1.1085.
- [18] F. Reynaldi Valerian, M. Syarief, and D. Abdul Fatah, "Klasifikasi Tingkat Obesitas Menggunakan Metode Gbm Dan Confusion Matrix," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 9, no. 2, pp. 2242–2249, 2025, doi: 10.36040/jati.v9i2.13062.