



Analisis Komparatif Sentimen Ulasan Google Play Store pada Shopee dan Tokopedia Menggunakan Algoritma Naive Bayes

Apria Andika¹, Hary Sabita²

Teknik Informatika, Ilmu Komputer, Institut Informatika dan Bisnis Darmajaya

¹apriaandikaa@gmail.com, ²hary.sabita@darmajaya.ac.id

Abstract

Shopee and Tokopedia are the two dominant e-commerce platforms in Indonesia. Both rely on consumer reviews from the Google Play Store as indicators of service success. A total of 60,000 reviews were collected between May 1 and June 18, 2026. To ensure computational efficiency without compromising statistical validity, the dataset was reduced to 500 samples per platform, calculated using Slovin's formula with a 5% margin of error. Text preprocessing was performed using Python, while modeling was conducted within an integrated environment using RapidMiner. The Naive Bayes algorithm was selected due to its low computational overhead and high accuracy in classifying large-scale text data. Cross-validation testing yielded accuracy rates of 78.20% for Shopee and 80.00% for Tokopedia. These results indicate that Tokopedia's positive sentiment (60.6%) surpassed that of Shopee (49.2%). Word Cloud visualizations highlighted distinct differences in reviews: Shopee's positive reviews were dominated by terms such as "barang" (item) and "mantap" (great), whereas Tokopedia's reviews featured prominent terms like "Indonesia" and "produk" (product). Regarding complaints, Shopee users focused on promotional aspects ("buruk"/bad, "diskon"/discount), while Tokopedia users cited technical issues ("gabisa"/unable, "lama"/slow, "ribet"/complicated). These findings demonstrate the effectiveness of the adopted methodology in mapping sentiment to provide strategic insights for e-commerce service improvement. Unlike previous studies that typically evaluated only a single platform, the novelty of this research lies in its direct comparative analysis of both platforms, reinforced by a hybrid computational scheme and visual extraction.

Keywords: Sentiment Analysis, Text Mining, Naive Bayes, Shopee, Tokopedia.

Abstrak

Shopee dan Tokopedia merupakan dua platform niaga elektronik (*e-commerce*) yang paling mendominasi di Indonesia. Keduanya mengandalkan ulasan konsumen di Google Play Store sebagai indikator keberhasilan layanan. Sebanyak 60.000 data ulasan dikumpulkan selama periode 1 Mei hingga 18 Juni 2026. Untuk efisiensi komputasi tanpa mengurangi validitas statistik, data direduksi menjadi 500 sampel per platform yang dihitung menggunakan pendekatan rumus Slovin dengan margin kesalahan 5%. Tahap prapemrosesan teks dikerjakan menggunakan Python, sementara pemodelan dilakukan secara terintegrasi menggunakan RapidMiner. Algoritma *Naive Bayes* dipilih karena komputasinya yang ringan namun memiliki akurasi tinggi dalam mengklasifikasikan data teks berskala besar. Pengujian *Cross-Validation* menghasilkan akurasi sebesar 78,20% untuk Shopee dan 80,00% untuk Tokopedia. Hasil ini menunjukkan bahwa sentimen positif Tokopedia (60,6%) lebih unggul dibandingkan dengan Shopee (49,2%). Visualisasi *Word Cloud* mempertegas perbedaan ulasan; ulasan positif Shopee didominasi kata "barang" dan "mantap", sedangkan Tokopedia menonjol pada kata "Indonesia" dan "produk". Pada sisi keluhan, pengguna Shopee menyoroti aspek promosi ("buruk", "diskon"), sementara pengguna Tokopedia mengeluhkan kendala teknis ("gabisa", "lama", "ribet"). Hasil ini membuktikan bahwa pendekatan metodologi tersebut efektif dalam memetakan sentimen guna memberikan wawasan strategis bagi perbaikan layanan *e-commerce*. Berbeda dengan penelitian sebelumnya yang umumnya hanya mengevaluasi satu platform, kebaruan (*novelty*) dari penelitian ini secara tegas terletak pada analisis komparatif langsung antara kedua platform yang diperkuat dengan penggunaan skema komputasi hibrida serta ekstraksi visual.

Kata kunci: Analisis Sentimen, Text Mining, Naive Bayes, Shopee, Tokopedia.

© 2026 Author

Creative Commons Attribution 4.0 International License



1. Pendahuluan

Lonjakan aktivitas digital di Indonesia memicu pertumbuhan pesat pada sektor perdagangan elektronik (*e-commerce*)[1]. Saat ini, *e-commerce* didominasi oleh dua platform utama, yakni Shopee dan Tokopedia[2]. Loyalitas konsumen mereka tidak lagi hanya dibatasi oleh kelengkapan barang dan harga, tetapi juga sangat bergantung pada kenyamanan antarmuka (*interface*) serta performa pada layanan aplikasinya[3]. Kepuasan konsumen terhadap aplikasi menjadi indikator vital bagi keberlanjutan platform tersebut[4].

Salah satu cara untuk mengukur kepuasan konsumen terhadap aplikasi adalah melalui rekam jejak ulasan[5]. Ulasan ditulis oleh konsumen di layanan *Google Play Store* – platform distribusi digital. Platform distribusi ini menyimpan pandangan subjektif konsumen, seperti pujian atas kemudahan transaksi dan laporan keluhan teknis. Kendati demikian, ulasan nantinya akan terkumpul dan membentuk data yang tidak terstruktur dalam jumlah masif.

Data yang telah terkumpul selanjutnya bisa digunakan untuk memperoleh informasi tingkat kepuasan konsumen. Karena jumlahnya yang masif – dalam jumlah ribuan –, maka tidak mungkin memperoleh informasi dengan menggunakan proses analisis yang manual – data diperiksa satu per satu. Salah satu solusi untuk mengatasi masalah ini adalah menggunakan metode komputasi berbasis *text mining*.

Text mining adalah sebuah metode yang digunakan untuk mengolah data yang berbentuk teks[6]. Teks dikumpulkan kemudian dikelompokkan ke dalam sentimen positif maupun negatif. Umumnya, para peneliti menggunakan sebuah algoritma untuk mendukung metode ini, yaitu Naïve Bayes (NB)[7]. NB bekerja untuk mengklasifikasi sentimen dengan cara menghitung peluang munculnya kata dalam kumpulan teks. NB dipilih karena tingkat akurasi yang tinggi dan ringan untuk mengolah kumpulan teks yang berskala besar[8].

Penelitian sebelumnya telah membuktikan keandalan pendekatan algoritma NB dalam mengevaluasi ulasan pengguna aplikasi. Studi yang dilakukan oleh Isnain, Nofitasari, dan Kesuma (2021) menunjukkan bahwa implementasi algoritma NB sangat akurat dalam menakar sentimen kepuasan pada aplikasi Shopee. Di objek penelitian yang berbeda, Muntashir dan Bachtiar (2021) juga sukses menerapkan NB untuk

mendeteksi rasio keluhan dan apresiasi pada aplikasi Tokopedia. Walaupun riset telah banyak dilakukan, namun hanya fokus pada satu platform – Shopee atau Tokopedia. Hal ini menjadi sebuah kelemahan ketika ingin mengetahui perbandingan evaluasi ulasan kedua platform tersebut.

Untuk melengkapi kekurangan penelitian sebelumnya, penelitian ini menawarkan 3 hal baru, yaitu 1) membandingkan ulasan dua platform *e-commerce* terbesar, yaitu Shopee dan Tokopedia, 2) menggabungkan dua alat bantu (*tools*), yaitu *Python* dan *RapidMiner*, dan 3) menggunakan visualisasi awan kata atau *Word Cloud* untuk melihat langsung apa yang paling sering dibicarakan konsumen. Dengan cara ini, akar masalah dari setiap aplikasi pada Shopee dan Tokopedia bisa ditemukan dengan jelas, sehingga saran perbaikan yang diberikan benar-benar tepat sasaran dan sesuai dengan keluhan konsumen.

2. Metode Penelitian

Metodologi dalam penelitian dibagi menjadi 4 tahap, yaitu: 1) pengumpulan data, 2) prapemrosesan data, 3) pemodelan klasifikasi serta visualisasi, dan 4) evaluasi hasil[9]. Tahap 1 dan 2 dikerjakan menggunakan *Python*. *Python* digunakan untuk ekstraksi dan pembersihan data. Sementara itu, tahap 3 dan 4 diolah menggunakan *RapidMiner*. *RapidMiner* bekerja untuk membangun model prediksi dan menghasilkan visualisasi yang representatif[10]. Gambar 1 memperlihatkan tahapan-tahapan penelitian.



Gambar 1. Tahapan-tahapan Penelitian

2.1. Pengumpulan Data (Tahap 1)

Tahap 1 melakukan ekstraksi data dalam jumlah masif secara otomatis. Data yang diekstraksi adalah ulasan konsumen Shopee dan Tokopedia di *Google Play Store*. Proses ini dilakukan dalam rentang waktu antara 1 Mei hingga 18 Juni 2026 dan menghasilkan data mentah sebanyak 60.000 baris dalam format *.xlsx*. Data mentah berisi 10 atribut (Tabel 1).

Tabel 1. Atribut Data Mentah

No	Nama Atribut
1	<i>Review Id</i>
2	<i>Username</i>
3	<i>User Image</i>
4	<i>Content</i>
5	<i>Score</i>

6	<i>Thums up Count</i>
7	<i>Review Created Version</i>
8	<i>At</i>
9	<i>Reply Content</i>
10	<i>App Version</i>

Seluruh atribut diklasifikasikan dan dieliminasi agar menghasilkan ulasan yang memiliki 3 label, yaitu: 1) positif, 2) netral, dan 3) negatif. Proses pelabelan kelas sentimen pada penelitian ini dilakukan secara otomatis berdasarkan skor *rating* yang diberikan pengguna pada Google Play Store. Ulasan dengan *rating* 4 dan 5 dikategorikan sebagai sentimen positif, sedangkan ulasan dengan *rating* 1 dan 2 dilabeli sebagai sentimen negatif. Ulasan bernilai bintang 3 dieliminasi dari *dataset* karena dianggap netral dan berpotensi memunculkan ambiguitas pada pemodelan klasifikasi biner. Proses ini dilakukan dengan menambahkan satu atribut baru, yaitu *Sentiment* – atribut nomor 11. Tahap 1 diakhiri dengan menerapkan teknik pengambilan sampel acak. Teknik ini dilakukan dengan menambah atribut bantuan – atribut nomor 12 – untuk menghasilkan angka. Angka tersebut diletakkan pada atribut nomor 12 dan dikunci menjadi nilai permanen. Penambahan atribut pada data mentah ditunjukkan pada Tabel 2.

Tabel 2. Atribut Tambahan

No	Nama Atribut
11	<i>Sentiment</i>
12	<i>Cleaned Content</i>

Setelah nilainya tetap, seluruh baris data diurutkan berdasarkan besaran angka acak tersebut. Dari hasil pengurutan ini, dihasilkan sebuah *dataset* yang berisi 500 baris data teratas yang mewakili masing-masing platform – Shopee dan Tokopedia. Pengambilan 500 sampel ulasan per platform tidak dilakukan secara acak tanpa dasar, melainkan dihitung menggunakan pendekatan statistik Slovin untuk memastikan tingkat keterwakilan data [11]. Dengan populasi (N) sebanyak 60.000 ulasan dan batas toleransi kesalahan (e) sebesar 5%, batas minimal sampel (n) dihitung menggunakan formula 1:

$$n = \frac{N}{1+N(e)^2} \quad (1)$$

Berdasarkan perhitungan tersebut, jumlah sampel minimal yang dibutuhkan adalah 397 data. Oleh karena itu, penetapan 500 sampel ulasan melalui teknik *random sampling* pada penelitian ini dipastikan telah memenuhi dan melampaui syarat validitas statistik.

2.2. Prapemrosesan Data (Tahap 2)

Tahap 2 berfokus pada atribut nomor 4 – *content* – yang berisi ulasan-ulasan. Kumpulan ulasan dipisahkan dari derau—informasi pada teks yang tidak relevan. Proses pemisahan dilakukan menggunakan *Python* melalui *Google Colab* – platform berbasis *cloud* untuk menulis program.

Tahap 2 terdiri dari 5 langkah utama yang dieksekusi secara berurutan, yakni: 1) *case folding*, 2) *cleansing*, 3) *tokenizing*, 4) *stopword removal*, dan 5) *normalization* dan *stemming*.

Pada langkah 1, dilakukan proses pengubahan seluruh ulasan teks menjadi huruf kecil. Selanjutnya, langkah 2 membersihkan ulasan dari entitas yang tidak relevan, seperti angka, tanda baca, tautan web, dan emotikon. Teks yang sudah bersih tersebut kemudian dipecah menjadi potongan kata tunggal melalui langkah 3. Setelah itu, langkah 4 mengeksekusi penyaringan untuk membuang kata-kata umum atau kata hubung (seperti "dan", "yang", "di") yang sering muncul namun tidak memiliki bobot sentimen. Langkah terakhir, langkah 5, berfungsi untuk memperbaiki kosakata tidak baku menjadi bahasa standar, sekaligus mengembalikan kata-kata berimbuhan ke bentuk dasarnya.

Hasil akhir dari tahap 2 adalah sebuah *dataset*. *Dataset* ini hanya berisi tiga atribut yang telah bersih, yaitu atribut nomor 4, 11, dan 12 (lihat Tabel 1 dan 2).

2.3. Pemodelan Klasifikasi dan Visualisasi (Tahap 3)
Tahap 3 dilakukan untuk membuat model prediksi dan visualisasi *dataset*. Tahap ini dirancang dengan menyusun beberapa operator pada *RapidMiner* secara berurutan (Tabel 3).

Tabel 3. Operator RapidMiner

No	Nama Operator
1	<i>Read Excel</i>
2	<i>Select Attributes</i>
3	<i>Set Role</i>
4	<i>Nominal to Text</i>
5	<i>Process Document from Data</i>
6	<i>Cross Validation</i>
7	<i>Performance</i>

Operator 1 berfungsi untuk memasukkan *dataset* ke *RapidMiner*. Selanjutnya, operator 2 menyortir dua atribut, yaitu atribut 11 dan 12 (lihat Tabel 2). Langkah berikutnya menggunakan operator 3 untuk menetapkan atribut 11 sebagai label yang akan diprediksi. Setelah itu, operator 4 mengubah *filter type* pada atribut 12 menjadi bertipe tunggal. Hal ini bertujuan agar atribut 12 berformat teks—yang sebelumnya terbaca oleh sistem sebagai format kategori.

Atribut 12 yang sudah berformat teks kemudian dikonversi menjadi representasi angka menggunakan operator 5 melalui teknik perhitungan *Term Frequency-Inverse Document Frequency* (TF-IDF) [12]. TF-IDF menilai seberapa penting sebuah kata berdasarkan frekuensinya di dalam keseluruhan ulasan.

Proses evaluasi dan klasifikasi dieksekusi melalui operator 6. Operator ini berfungsi membagi data pelatihan dan pengujian dengan menerapkan teknik pengujian *10-fold Cross-Validation*, di mana *dataset* dibagi menjadi 10 bagian (9 bagian untuk data latih

dan 1 bagian untuk data uji secara bergantian). Operator 6 memiliki 2 tahapan subproses utama. Pertama, algoritma NB digunakan untuk mengenali pola kata yang termasuk ke dalam kelas ulasan positif atau negatif. Konfigurasi model NB menggunakan parameter bawaan (*default*) dengan mengaktifkan fitur *Laplace Smoothing* untuk mencegah probabilitas nol pada kata yang tidak dikenali[13]. Hal ini dilakukan misalnya dengan mencatat secara statistik bahwa kata "lemot" atau "kecewa" sangat identik dengan ulasan negatif, sedangkan kata "mudah" atau "membantu" mencerminkan ulasan positif. Kedua, *Apply Model* berfungsi untuk menguji ketepatan algoritma NB dalam memprediksi label pada atribut 11. Rangkaian evaluasi ditutup dengan operator 7. Operator ini mengukur nilai akhir untuk menghitung persentase tingkat akurasi algoritma NB dalam mengklasifikasikan sentimen secara keseluruhan.

Setelah pemodelan klasifikasi selesai, proses difokuskan pada ekstraksi visual. Dataset divisualisasikan ke dalam bentuk grafik *Word Cloud*. Representasi grafik ini bertujuan untuk memetakan dan menampilkan secara grafis kata-kata yang paling dominan pada ulasan pengguna aplikasi Shopee dan Tokopedia. Semakin besar dan tebal ukuran huruf sebuah kata, semakin tinggi pula frekuensi kemunculannya di dalam ulasan, sehingga keluhan atau pujian utama dari konsumen dapat diidentifikasi dengan mudah.

2.4. Evaluasi Hasil (Tahap 4)

Evaluasi dilakukan untuk mengukur seberapa baik dan akurat algoritma NB dalam mengklasifikasikan sentimen ulasan. Alat ukur yang digunakan pada tahap ini adalah *Confusion Matrix* (CM). CM adalah sebuah tabel yang digunakan untuk membandingkan data aktual dengan hasil prediksi. Nilai prediksi ditentukan berdasarkan akurasi dari empat indikator seperti pada formula 2 berikut:

$$\text{Akurasi} = \frac{(TP+TN)}{(TP+TN+FP+FN)} \times 100\% \quad (2)$$

Akurasi adalah rasio persentase yang menunjukkan total prediksi benar, baik positif maupun negatif, berbanding dengan keseluruhan data ulasan yang diuji dengan empat indikator (Tabel 4).

Tabel 4. Indikator Confusion Matrix

No	Indikator
1	True Positive (TP)
2	True Negative (TN)
3	False Positive (FP)
4	False Negative (FN)

TP merepresentasikan jumlah ulasan positif yang berhasil diprediksi dengan benar, dan TN sebagai jumlah ulasan negatif yang berhasil diprediksi. Selain itu, terdapat indikator kesalahan prediksi berupa FP, yaitu kondisi di mana ulasan yang sebenarnya negatif justru diprediksi sebagai ulasan positif, serta FN, di

mana ulasan positif disalahartikan sebagai ulasan negatif oleh algoritma NB.

Selain mengukur tingkat akurasi (rasio prediksi benar terhadap keseluruhan data), evaluasi klasifikasi pada penelitian ini juga menyajikan pengukuran kinerja yang lebih komprehensif melalui matriks *Precision*, *Recall*, dan *F1-Score*. *Precision* digunakan untuk mengukur tingkat ketepatan prediksi kelas positif, *Recall* untuk mengukur sensitivitas model dalam mengenali ulasan kelas positif yang sebenarnya, dan *F1-Score* sebagai rata-rata harmonis antara keduanya. Keseluruhan indikator tersebut dihitung menggunakan formula evaluasi standar berdasarkan nilai yang dihasilkan oleh *Confusion Matrix* berikut ini:

$$\text{Precision} = \frac{TP}{TP+FP} \times 100\% \quad (3)$$

$$\text{Recall} = \frac{TP}{TP+FN} \times 100\% \quad (4)$$

$$F1 - \text{Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (5)$$

Semakin tinggi persentase akurasi beserta nilai evaluasi matriks lainnya (*Precision*, *Recall*, dan *F1-Score*), semakin layak dan valid model klasifikasi yang dibangun dengan algoritma NB untuk menarik kesimpulan dari ulasan konsumen Shopee dan Tokopedia.

Selain mengevaluasi performa algoritma, penelitian ini juga melakukan uji signifikansi statistik untuk memastikan bahwa perbedaan sentimen antara Shopee dan Tokopedia benar-benar bermakna dan bukan sekadar faktor kebetulan (*sampling error*). Pengujian dilakukan menggunakan metode *Chi-Square* (χ^2) dengan tingkat signifikansi (α) sebesar 5% atau 0,05. Formula 6 perhitungan *Chi-Square* adalah sebagai berikut:

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} \quad (6)$$

Di mana O_i adalah nilai frekuensi yang diobservasi (*observed frequency*) dan E_i adalah nilai frekuensi harapan (*expected frequency*). Nilai *Chi-Square* hitung kemudian dibandingkan dengan nilai kritis *Chi-Square* pada tabel distribusi dengan derajat kebebasan (*degree of freedom / df*) = 1. Jika nilai χ^2 hitung lebih besar daripada nilai kritis, maka Hipotesis Nol (H_0) ditolak, yang menunjukkan adanya perbedaan yang signifikan secara statistik antara kedua kelompok data tersebut.

3. Hasil dan Pembahasan

RapidMiner menghasilkan akurasi model dan visualisasi *word cloud*. Akurasi menyajikan rasio angka kepuasan konsumen; analisis ini juga mengungkap alasan spesifik—baik berupa apresiasi layanan maupun keluhan teknis—yang paling

memengaruhi kenyamanan pengguna saat bertransaksi di kedua platform *e-commerce* tersebut.

3.1. Klasifikasi Sentimen

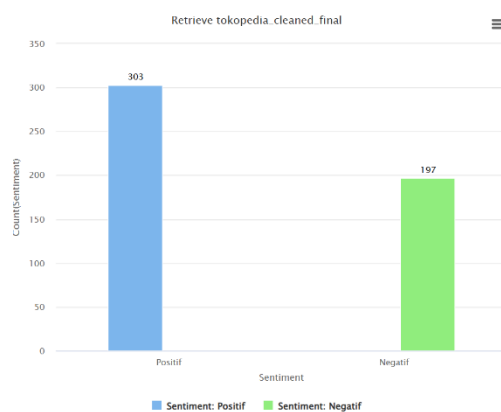
Pengklasifikasian dilakukan terhadap 1.000 sampel ulasan (500 data per platform) untuk memetakan ulasan secara biner. Pemodelan sentimen menunjukkan adanya perbedaan tingkat kepuasan yang cukup jelas antara pengguna Shopee dan Tokopedia. Komparasi sentimen dari kedua platform tersebut disajikan secara rinci pada Tabel 5.

Tabel 5. Komparasi Sentimen

E-commerce	Sentimen Positif	Sentimen Negatif	Total Data
Shopee	246 (49,2%)	254 (50,8%)	500
Tokopedia	303 (60,6%)	197 (39,4%)	500



Gambar 2. Sentimen Positif dan Negatif Shopee.



Gambar 3. Sentimen Positif dan Negatif Tokopedia.

Berdasarkan Tabel 5, pengguna Tokopedia menunjukkan kecenderungan tingkat kepuasan yang lebih tinggi dengan persentase sentimen positif mencapai 60,6%. Sebaliknya, ulasan publik terhadap layanan Shopee cenderung lebih berimbang, namun sentimen negatif sedikit mendominasi di angka 50,8%. Hal ini mengindikasikan bahwa secara akumulatif, antarmuka dan layanan Tokopedia pada

periode pengamatan dinilai lebih stabil oleh penggunanya.

Untuk membuktikan apakah perbedaan sentimen antara kedua platform tersebut bermakna secara statistik, penelitian ini melakukan uji signifikansi menggunakan Uji *Chi-Square* (χ^2) [14]. Berdasarkan jumlah sampel ($N = 1.000$) dengan distribusi sentimen Shopee (246 Positif, 254 Negatif) dan Tokopedia (303 Positif, 197 Negatif), hasil perhitungan menghasilkan nilai χ^2 sebesar 13,12. Mengingat nilai kritis χ^2 untuk derajat kebebasan ($df = 1$) pada tingkat signifikansi 5% ($\alpha = 0,05$) adalah 3,841, maka nilai yang diperoleh ($13,12 > 3,841$) menunjukkan $p\text{-value} < 0,05$. Hal ini membuktikan bahwa Hipotesis Nol (H_0) ditolak [15]. Dengan kata lain, terdapat perbedaan sentimen yang signifikan secara statistik antara pengguna Shopee dan Tokopedia, di mana tingkat kepuasan pengguna Tokopedia secara meyakinkan terbukti lebih tinggi pada periode pengamatan ini.

3.2. Evaluasi Performa Algoritma *Naive Bayes* (NB)

Untuk memvalidasi keandalan klasifikasi pada Tabel 5, performa algoritma NB dievaluasi menggunakan pengujian *Cross-Validation* dan CM. Pengujian ini mengukur sejauh mana sistem mampu memprediksi kelas sentimen dengan tepat dibandingkan dengan data aktual [16]. Hasil komputasi di *RapidMiner* mencatatkan nilai akurasi yang sangat baik. Perbandingan CM dari kedua *e-commerce* disajikan pada Tabel 6.

Tabel 6. Perbandingan Nilai Indikator *Confusion Matrix* Shopee dan Tokopedia

Matriks Evaluasi	Nilai Shopee	Nilai Tokopedia
TP	192	242
TN	199	158
FP	55	39
FN	54	61
Total Data	500	500

Berdasarkan data pada Tabel 6, tingkat akurasi dihitung dengan membagi total prediksi sentimen yang benar (TP dan TN) dengan keseluruhan data uji. Perhitungan menggunakan formula akurasi (Formula 1) menunjukkan bahwa model klasifikasi untuk Shopee mencapai tingkat akurasi sebesar 78,20%. Sementara itu, performa model pada Tokopedia menunjukkan hasil yang sedikit lebih unggul di angka 80,00%.

Selain akurasi, performa algoritma juga dievaluasi menggunakan matriks *Precision*, *Recall*, dan *F1-Score* untuk mengukur tingkat presisi dan sensitivitas pengenalan ulasan positif dan negatif [17]. Komparasi keseluruhan hasil evaluasi performa dari kedua platform tersebut disajikan pada Tabel 7.

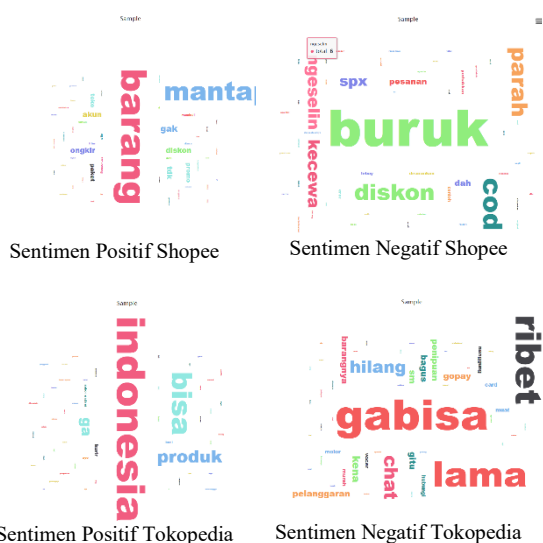
Matriks Evaluasi	Kinerja Model Shopee	Kinerja Model Tokopedia
Accuracy	78,20%	80,00%
Precision	77,73%	86,12%
Recall	78,05%	79,87%
F1-Score	77,89%	82,88%

3.3. Analisis Ulasan Pengguna Melalui Word Cloud
Langkah berikutnya adalah menggali lebih dalam ulasan konsumen menggunakan visualisasi *Word Cloud*. Teknik visualisasi ini dipilih karena kemampuannya dalam menyoroti kata-kata yang paling sering muncul melalui ukuran teks yang proporsional, sehingga sangat memudahkan identifikasi topik atau akar permasalahan utama dari kumpulan data teks yang besar [19]. Kumpulan ulasan yang menunjukkan sentimen positif dan negatif disajikan secara jelas pada Gambar 4 berikut.

Dari hasil analisis kata ini, terlihat bahwa meskipun Shopee dan Tokopedia bergerak di bidang industri yang sama, tantangan yang mereka hadapi ternyata sangat berbeda. Shopee perlu mengevaluasi kembali aturan diskon, sistem pembayaran di tempat atau *Cash on Delivery* (COD), dan tampilan iklannya. Sementara itu, Tokopedia harus segera meningkatkan kecepatan dan kelancaran aplikasinya agar pengguna tidak lagi merasa kesulitan saat bertransaksi. Rekomendasi strategis berbasis pemodelan data semacam ini terbukti krusial dan dapat dijadikan landasan perbaikan antarmuka bagi layanan bisnis *e-commerce* yang berkelanjutan[20].

Berdasarkan hasil penelitian yang telah dilakukan, algoritma NB terbukti akurat dalam memilah ribuan ulasan konsumen. Sistem ini mampu mengenali ulasan pengguna dengan tingkat keberhasilan 80,00% pada Tokopedia dan 78,20% pada Shopee. Melalui perhitungan tersebut, ditemukan fakta bahwa Tokopedia mencatatkan kinerja yang lebih baik di mata konsumen dengan perolehan 60,6% ulasan bernada positif. Di sisi lain, respons konsumen terhadap Shopee menunjukkan hasil yang lebih berimbang, meskipun keluhan pengguna sedikit lebih mendominasi pada angka 50,8%.

Penelitian selanjutnya, disarankan untuk tidak hanya mengandalkan satu algoritma, tetapi



Berdasarkan Gambar 4, kita bisa melihat pengalaman pengguna dari Shopee dan Tokopedia dengan lebih jelas. Pada ulasan positif, pengguna Shopee lebih banyak memuji kualitas barang yang mereka beli. Hal ini terlihat dari seringnya munculnya kata "barang" dan "mantap". Sementara itu, ulasan positif di Tokopedia menunjukkan rasa bangga terhadap produk lokal dengan seringnya munculnya kata "Indonesia", yang disertai pujian pada layanan melalui kata "produk", "bisa", dan "bagus".

membandingkannya dengan algoritma lain agar hasil analisis menjadi lebih tajam dan akurat. Selain itu, cakupan data ulasan juga bisa diperluas, tidak hanya dari aplikasi saja, tetapi juga dari media sosial. Dengan cara ini, hasil penelitian akan lebih kaya dan benar-benar mencerminkan pendapat luas dari masyarakat sebagai konsumen.

Daftar Rujukan

- [1] E. Y. Nasution, P. Hariani, L. S. Hasibuan, and W. Pradita, "Perkembangan Transaksi Bisnis E-Commerce terhadap Pertumbuhan Ekonomi di Indonesia," *Jesya*, vol. 3, no. 2, pp. 506–519, 2020, doi: 10.36778/jesya.v3i2.227.
- [2] D. Susanti, "Analisis Perbandingan Marketplace Shopee Dan Tokopedia Terhadap Minat Beli Konsumen Digital," vol. 2, no. 1, pp. 1084–1093, 2026.
- [3] M. H. Hamdanuddinsyah, M. Hanafi, and P. Sukmasetya, "Perancangan UI/UX Aplikasi Buku Online Mizanstore Berbasis Mobile Menggunakan User Centered Design," *J. Inf. Syst. Res.*, vol. 4, no. 4, pp. 1464–1475, 2023, doi: 10.47065/josh.v4i4.3850.
- [4] Taqdirul Alim, D. Zufriah, and Muhammad Isnaini Fathoni, "Pengaruh Kepuasan Pelanggan dan Kualitas Layanan Terhadap Loyalitas Pelanggan pada E-Commerce Shopee," *J. Bisnis, Manajemen, dan Akunt.*, vol. 12, no. 1, pp. 66–74, 2025, doi: 10.54131/jbma.v12i1.215.
- [5] Syafrizal, M. Afdal, and R. Novita, "Sentiment Analysis of PLN Mobile Application Review Using Naïve Bayes Classifier and K-Nearest Neighbor Algorithm," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 4, no. January, pp. 10–19, 2024.
- [6] S. Sukriadi, I. Ismail, and A. M. Andzar, "Penerapan Text Mining Dalam Klasifikasi Judul Skripsi Yang Diusulkan Mahasiswa Menggunakan Metode Naïve Bayes," *J. Ilm. Sist. Inf. dan Tek. Inform.*, vol. 6, no. 2, pp. 184–196, 2023, doi: 10.57093/jisti.v6i2.174.
- [7] A. Rahman, E. Utami, and S. Sudarmawan, "Sentimen Analisis Terhadap Aplikasi pada Google Playstore Menggunakan Algoritma Naïve Bayes dan Algoritma Genetika," *J. Komtika (Komputasi dan Inform.)*, vol. 5, no. 1, pp. 60–71, 2021, doi: 10.31603/komtika.v5i1.5188.
- [8] A. Khoerunnisa, M. Fakhru Shiddiq, and S. Nurhayati, "Perbandingan Algoritma Naïve Bayes dan SVM dalam Analisis Sentimen Media Sosial di Indonesia," *J. Sist. Cerdas dan Komputasi*, vol. 10, no. 1, pp. 45–56, 2025.
- [9] V. Agustina and A. Herliana, "Analisis Sentimen Publik atas Kebijakan Efisiensi Anggaran 2025 dengan Text Mining dan Natural Language Processing," *J. Media Inform.*, vol. 6, no. 3, pp. 2182–2194, 2025, [Online]. Available: <https://ejournal.sisfokomtek.org/index.php/jumin/article/view/6301>
- [10] G. Shalihah, R. Kurniawan, and T. Suprpti, "Analisis Sentimen Ulasan Pelanggan Mie Gacoan Menggunakan Metode Naïve Bayes Classifier," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 8, no. 1, pp. 593–601, 2024, doi: 10.36040/jati.v8i1.8302.
- [11] B. Artono, "Jurnal Multidisiplin Sosiasl Humaniora," *J. Multidisiplin Sos. Hum.*, vol. 1, pp. 53–63, 24AD, [Online]. Available: <https://jurnal.ananpublisher.com/index.php/jmsh>
- [12] N. G. Yudiarta, M. Sudarma, and W. G. Ariastina, "Penerapan Metode Clustering Text Mining Untuk Pengelompokan Berita Pada Unstructured Textual Data," *Maj. Ilm. Teknol. Elektro*, vol. 17, no. 3, p. 339, 2018, doi: 10.24843/mite.2018.v17i03.p06.
- [13] Y. A. Singgalen, "Analisis Sentimen Wisatawan Melalui Data Ulasan Candi Borobudur di Tripadvisor Menggunakan Algoritma Naïve Bayes Classifier," *Build. Informatics, Technol. Sci.*, vol. 4, no. 3, 2022, doi: 10.47065/bits.v4i3.2486.
- [14] Y. H. D. Wangsajaya, E. Setyowati, and A. Wibowo, "Deteksi Emosi Teks X Berbahasa Indonesia Menggunakan Bi-LSTM dengan Seleksi Fitur Chi-Square," *J. Algoritm.*, vol. 22, no. 2, pp. 546–555, 2025, doi: 10.33364/algoritma/v.22-2.2658.
- [15] I. Nugraha, "Analisis Pengaruh Produk HP Samsung Terhadap Minat Beli Konsumen Dengan Menggunakan Metode Uji Chi-square," *J. TRINISTIK J. Tek. Ind. Bisnis Digit. dan Tek. Logistik*, vol. 3, no. 1, pp. 46–52, 2024, doi: 10.20895/trinistik.v3i1.1423.
- [16] K. Kenny, D. Arisandi, and T. Sutrisno, "Evaluasi Penilaian Kinerja Karyawan dengan Metode Naïve Bayes," *Comput. J. Comput. Sci. Inf. Syst.*, vol. 8, no. 1, pp. 110–118, 2024, doi: 10.24912/computatio.v8i1.20287.
- [17] Muhamad Fadli and Rizal Adi Saputra, "Klasifikasi dan Evaluasi Performa Model Random Forest untuk Prediksi Stroke," *JT J. Tek.*, vol. 12, no. 2, pp. 72–80, 2023, [Online]. Available: <http://jurnal.umt.ac.id/index.php/jt/index>
- [18] R. S. Nuraini *et al.*, "ANALISIS SENTIMEN PENGGUNA APLIKASI AGODA DI GOOGLE PLAY Sistem Informasi Universitas Amikom Purwokerto Abstraksi Keywords: Pendahuluan Tinjauan Pustaka Metode Penelitian," *J. Inf. Syst. Manag.*, vol. 7, no. 1, pp. 24–29, 2025.
- [19] R. Z. Firdaus, S. H. Wijoyo, and W. Purnomo, "Analisis Sentimen Berbasis Aspek Ulasan Pengguna Aplikasi Alfagift Menggunakan Metode Random Forest dan LDA," *J-Ptiik Ub*, vol. 9, no. 2, pp. 1–10, 2024, [Online]. Available: <https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/14482>
- [20] T. Tukino, "Penerapan Algoritma Convolutional Neural Network Untuk Klasifikasi Sentimen Pada Layanan e-Commerce," *J. Desain Dan Anal. Teknol.*, vol. 4, no. 1, pp. 44–53, 2025, doi: 10.58520/jddat.v4i1.72.