



Analisis Sentimen Ulasan Produk Shopee: Perbandingan Bag of Words dan TF-IDF

Hadit Halim Arshidik¹, Siti Hafizhah², Mokh. Fadli Rinjani³, Aldi Winata⁴

^{1,2,3,4} Universitas Lancang Kuning, Indonesia

***Correspondence:** hadit1905@gmail.com

Abstract

Perkembangan platform e-commerce mendorong meningkatnya jumlah ulasan pelanggan yang bersifat tekstual dan sulit dianalisis secara manual. Ulasan produk pada platform Shopee mengandung informasi penting terkait tingkat kepuasan pelanggan serta kualitas produk dan layanan. Penelitian ini bertujuan untuk menganalisis sentimen ulasan produk Shopee menggunakan pendekatan text mining dan klasifikasi teks. Data ulasan diperoleh dari dataset sekunder dan dilabeli berdasarkan nilai rating menjadi sentimen positif dan negatif. Tahapan penelitian meliputi text preprocessing, ekstraksi fitur teks, klasifikasi sentimen menggunakan beberapa algoritma, serta evaluasi kinerja model. Hasil penelitian menunjukkan bahwa pendekatan klasifikasi teks mampu mengklasifikasikan sentimen ulasan dengan akurasi yang baik, di mana beberapa model menunjukkan performa yang lebih seimbang dalam mendeteksi sentimen positif dan negatif. Selain itu, visualisasi data membantu dalam memahami pola dan karakteristik ulasan pelanggan secara lebih informatif.

Keywords: Analisis sentimen, ulasan produk, text mining, klasifikasi teks, e-commerce

Received: 01-06-2026 | Revised: 05-06-2026 | Accepted: 26-06-2026

Jstbs is licensed under a Creative Commons Attribution-Share Alike 4.0 International License



1. Pendahuluan

Perkembangan teknologi informasi dan internet telah mendorong pertumbuhan e-commerce secara pesat di Indonesia. Masyarakat kini semakin terbiasa melakukan transaksi jual beli secara daring karena dinilai lebih praktis, cepat, dan efisien. Salah satu platform e-commerce yang sangat populer adalah Shopee, yang menyediakan berbagai produk serta fitur ulasan sebagai sarana bagi pelanggan untuk berbagi pengalaman setelah melakukan pembelian^[1].

Ulasan produk yang ditulis oleh pengguna mengandung informasi penting mengenai kualitas produk, pelayanan penjual, serta tingkat kepuasan pelanggan. Informasi ini tidak hanya bermanfaat bagi calon pembeli, tetapi juga bagi penjual dan pihak platform untuk meningkatkan kualitas layanan. Namun, jumlah ulasan yang terus meningkat setiap harinya menyebabkan data tersebut sulit dianalisis secara manual karena berbentuk teks tidak terstruktur dan berjumlah sangat besar^[2].

Untuk mengatasi permasalahan tersebut, diperlukan pendekatan otomatis yang mampu mengolah dan menganalisis data teks dalam jumlah besar. Salah satu metode yang banyak digunakan adalah analisis sentimen, yaitu teknik untuk mengidentifikasi kecenderungan opini atau perasaan seseorang terhadap suatu objek, baik bersifat positif

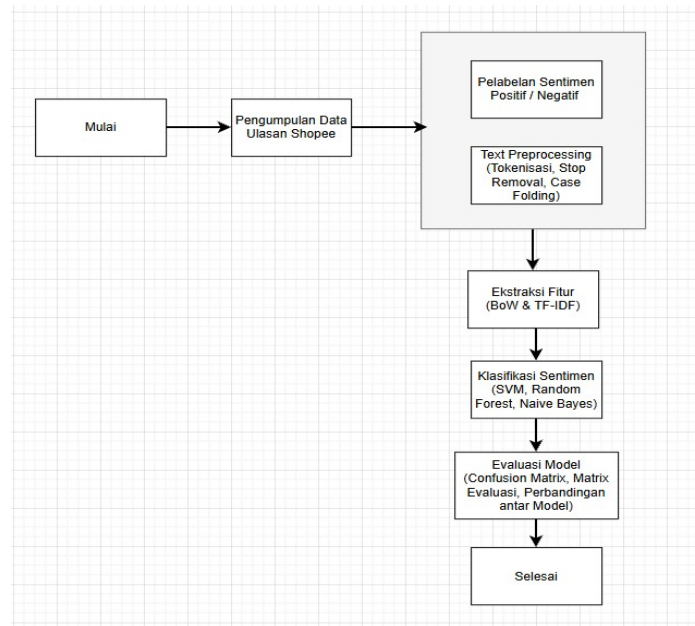
maupun negatif. Dengan analisis sentimen, ulasan pelanggan dapat diklasifikasikan secara sistematis sehingga pola kepuasan atau ketidakpuasan dapat diketahui dengan lebih cepat dan akurat[3].

Pendekatan text mining dan klasifikasi teks merupakan metode yang efektif dalam analisis sentimen karena mampu mengekstraksi informasi penting dari data teks. Proses ini melibatkan beberapa tahapan seperti pembersihan data, tokenisasi, ekstraksi fitur, serta penerapan algoritma klasifikasi untuk menentukan kategori sentimen. Dengan metode ini, data ulasan Shopee dapat diolah menjadi informasi yang lebih terstruktur dan mudah dianalisis[4].

Berdasarkan hal tersebut, penelitian ini bertujuan untuk menganalisis sentimen ulasan produk Shopee menggunakan pendekatan text mining dan klasifikasi teks. Selain mengklasifikasikan ulasan ke dalam sentimen positif dan negatif, penelitian ini juga menyajikan visualisasi data untuk membantu memahami distribusi sentimen dan karakteristik ulasan pelanggan. Diharapkan hasil penelitian ini dapat memberikan gambaran yang lebih jelas mengenai tingkat kepuasan pengguna Shopee serta mendukung pengambilan keputusan yang lebih baik bagi pihak terkait[5].

2. Metode

Tahapan penelitian ini menggambarkan secara menyeluruh alur kerja sistem analisis sentimen yang digunakan untuk mengklasifikasikan ulasan pengguna Shopee ke dalam dua kategori utama, yaitu sentimen positif dan sentimen negatif, dengan memanfaatkan pendekatan machine learning. Sistem ini dirancang untuk mengolah data ulasan dalam bentuk teks yang bersifat tidak terstruktur menjadi informasi yang bermakna sehingga dapat digunakan untuk menilai persepsi dan kepuasan pengguna terhadap produk maupun layanan yang tersedia di platform Shopee. Proses dimulai dari tahap pengumpulan data ulasan, di mana komentar pengguna yang mengandung opini dan pengalaman dikumpulkan sebagai bahan analisis. Data tersebut kemudian melalui tahap pelabelan sentimen untuk memberikan kelas yang jelas pada setiap ulasan, sehingga dapat digunakan sebagai data latih dalam proses pembelajaran mesin. Selanjutnya, data teks diproses melalui tahapan preprocessing seperti tokenisasi, case folding, dan penghapusan stopword guna membersihkan data dari elemen yang tidak relevan dan mengurangi noise. Setelah data bersih, dilakukan proses ekstraksi fitur menggunakan metode Bag of Words (BoW) dan Term Frequency-Inverse Document Frequency (TF-IDF) untuk mengubah data teks menjadi representasi numerik yang dapat dipahami oleh algoritma machine learning. Representasi ini menjadi dasar bagi model klasifikasi untuk mempelajari pola-pola yang membedakan ulasan positif dan negatif. Tahap berikutnya adalah proses klasifikasi sentimen menggunakan beberapa algoritma, yaitu Support Vector Machine (SVM), Random Forest, dan Naive Bayes. Penggunaan lebih dari satu algoritma bertujuan untuk membandingkan kinerja masing-masing model dalam mengklasifikasikan sentimen ulasan. Hasil klasifikasi kemudian dievaluasi menggunakan confusion matrix dan berbagai metrik evaluasi seperti akurasi, precision, recall, dan F1-score untuk mengetahui tingkat keakuratan dan keandalan model. Dengan adanya tahapan-tahapan tersebut, sistem ini mampu memberikan gambaran yang lebih objektif mengenai kecenderungan sentimen pengguna Shopee. Informasi yang dihasilkan dapat dimanfaatkan sebagai bahan evaluasi kualitas produk dan layanan, serta sebagai dasar pengambilan keputusan bagi pihak penjual maupun pengelola platform dalam meningkatkan kepuasan pelanggan.



Gambar 1. Tahapan Penelitian

2.1 Pengumpulan Data

Tahap pertama adalah mengumpulkan data berupa teks ulasan pelanggan Shopee. Data ini diperoleh dari hasil crawling atau dataset publik yang berisi komentar pengguna terhadap produk atau layanan. Setiap data direpresentasikan sebagai:

$$D = \{d_1, d_2, d_3, \dots, d_n\}$$

2.2 Pelabelan

Sebagai ulasan diberi label sentimen:

A. Positif (+1+1+1)

B. Negatif (-1-1-1)

Sehingga menjadi :

$$D = \{(d_1, y_1), (d_2, y_2), \dots, (d_n, y_n)\}$$

2.3 Text Preprocessing

Tahap ini bertujuan membersihkan data text agar siap di proses oleh sistem.

A. Tokenisasi

Memecah kalimat menjadi kata:

$$T(d_i) = \{w_1, w_2, w_3, \dots, w_m\}$$

B. Case Folding

Mengubah seluruh huruf menjadi kecil:

$$W = \text{lower}(w)$$

C. Stopword Removal

Menghapus kata umum seperti "dan", "yang", "di".

$$T^1 = T(d_i) - S$$

2.4 Ekstraksi Fitur

Pendekatan Bag of Words (BoW) merupakan metode ekstraksi fitur yang digunakan untuk mengklasifikasikan dokumen dengan cara mengevaluasi frekuensi kemunculan kata kata unik dalam kumpulan teks yang besar. BoW menawarkan kesederhanaan dan

fleksibilitas yang memungkinkan representasi teks berdasarkan pola kemunculan kata dalam suatu dokumen, sehingga cocok untuk berbagai aplikasi analisis teks. Bobot kata dihitung menggunakan Term Frequency (TF), Inverse Document Frequency (IDF), TF-IDF

$$\text{IDF} = \log N_{\text{dft}}$$

2.5 Klasifikasi Sentimen

A. Support Vector Machine

SVM mencari hyperplane optimal :

$$w \cdot x + b = 0$$

Dengan fungsi Klasifikasi:

$$(x) = \text{sign}(w \cdot x + b)$$

B. Naive Bayes

Menggunakan teorema bayes:

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

Kelas dipilih berdasarkan :

$$C = \text{argmax}_C P(C) \prod_{i=1}^n P(x_i | C)$$

C. Random Forest

Menggunakan banyak pohon keputusan :

$$RF(x) = T_1 t = 1 \sum T h t(x)$$

2.6 Rumus Perhitungan Model dan Evaluasi

Pada penelitian ini, proses klasifikasi sentimen dilakukan menggunakan beberapa algoritma machine learning. Untuk mendukung proses klasifikasi dan evaluasi model, digunakan beberapa rumus yang berkaitan dengan algoritma Support Vector Machine, Random Forest, serta metrik evaluasi model. Rumus-rumus tersebut digunakan untuk menjelaskan dasar matematis dari proses klasifikasi dan pengukuran performa model.

2.6.1 Rumus Support Vector Machine

Support Vector Machine bekerja dengan mencari hyperplane terbaik yang dapat memisahkan data ke dalam kelas sentimen positif dan negatif. Fungsi klasifikasi SVM dapat dituliskan sebagai berikut:

$$f(x) = w \cdot x + b$$

Keterangan:

$f(x)$ = fungsi klasifikasi

w = bobot fitur

x = vektor fitur

b = bias

Adapun batas pemisahan data pada SVM dinyatakan dengan persamaan sebagai berikut:

$$y_i(w \cdot x_i + b) \geq 1$$

Tujuan utama SVM adalah memaksimalkan margin dengan meminimalkan fungsi berikut:

$$\min \frac{1}{2} \|w\|^2$$

Berdasarkan persamaan tersebut, SVM berusaha memperoleh hyperplane yang memiliki margin paling optimal sehingga mampu membedakan kelas sentimen secara lebih akurat.

2.6.2 Rumus Gini Index Random Forest

Random Forest menggunakan banyak pohon keputusan dalam proses klasifikasi. Pada setiap pohon keputusan, pemilihan node terbaik dapat dilakukan menggunakan nilai Gini Index. Rumus Gini Index adalah sebagai berikut:

$$\text{Gini} = 1 - \sum p_i^2$$

Keterangan:

Gini = nilai ketidakmurnian data

p_i = proporsi data pada kelas ke- i

c = jumlah kelas

Semakin kecil nilai Gini Index, maka semakin baik pemisahan data yang dilakukan oleh pohon keputusan. Dengan demikian, Random Forest dapat menentukan cabang keputusan yang paling sesuai dalam membedakan kelas sentimen.

2.6.3 Rumus Evaluasi Model

Evaluasi model dilakukan menggunakan confusion matrix yang terdiri dari nilai True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN). Berdasarkan nilai tersebut, kinerja model dihitung menggunakan accuracy, precision, recall, dan F1-score.

A. Accuracy

$$\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{TN} + \text{FP} + \text{FN})$$

Accuracy mengukur proporsi seluruh prediksi yang benar terhadap total data yang diuji.

B. Precision

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP})$$

Precision mengukur ketepatan model dalam memprediksi data positif.

C. Recall

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN})$$

Recall mengukur kemampuan model dalam menemukan seluruh data positif yang sebenarnya.

D. F1-Score

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

F1-Score mengukur keseimbangan antara nilai precision dan recall.

Keterangan confusion matrix:

TP = data positif yang diprediksi benar sebagai positif

TN = data negatif yang diprediksi benar sebagai negatif

FP = data negatif yang salah diprediksi sebagai positif

FN = data positif yang salah diprediksi sebagai negatif

Dengan adanya metrik evaluasi tersebut, performa masing-masing algoritma klasifikasi dapat dianalisis secara lebih objektif, baik dari sisi ketepatan prediksi, kemampuan mendeteksi kelas positif, maupun keseimbangan antara precision dan recall.

2.7 Evaluasi

Setelah diketahui confusion matrixnya, nilai accuracy, precision, recall, dan F1-score dihitung berdasarkan persamaan yang telah dijelaskan pada Subbab 2.6. Hasil perhitungan tersebut digunakan untuk menilai tingkat keakuratan dan keandalan model klasifikasi.

- a. Akurasi
Mengukur proporsi prediksi yang benar terhadap total prediksi yang dibuat oleh model, dihitung menggunakan persamaan
- b. Precision
Mengukur seberapa tepat prediksi model dengan menghitung rasio prediksi positif yang benar terhadap semua prediksi positif, dihitung menggunakan persamaan.
- c. Recall
Mengukur kemampuan model untuk menemukan semua contoh positif yang ada dalam data, dihitung menggunakan persamaan.
- d. F1-Score
Perhitungan yang menggambarkan perbandingan antara precision dan recall. Jika nilai FN dan FP tidak men-dekati sebaiknya digunakan f1-score dibandingkan nilai accuracy, dihitung menggunakan persamaan.

3. Hasil dan Pembahasan

Pada tahap ini akan di tampilkan hasil yang di peroleh dari setiap proses yang telah di lakukan sebelumnya:

3.1 Pengumpulan Data

Penelitian ini menggunakan dataset sekunder berupa ulasan produk Shopee yang telah dipublikasikan secara terbuka pada Kaggle. Struktur dan isi dataset yang digunakan ditampilkan pada gambar berikut:

Table 1. Isi Dataset Ulasan Shopee

No	User Name	Score	Content
1	Pengguna google	1	Mohon maaf kenapa seiap kali ingin melihat ha..
2	Pengguna google	5	ulala
3	Pengguna google	5	Mantul
4	Pengguna google	5	Bagus
5	Pengguna google	3	Ini knp vidio shoopenya BLANK HITAM?

3.2 Pelabelan

Pelabelan merupakan tahap pemberian label pada data ulasan untuk menentukan kategori atau kelas tertentu yang akan digunakan dalam proses klasifikasi. Pada penelitian ini, pelabelan dilakukan berdasarkan nilai rating (score) yang diberikan oleh pengguna pada ulasan produk Shopee. Nilai rating tersebut digunakan sebagai acuan untuk mengelompokkan ulasan ke dalam kategori sentimen. Ulasan dengan nilai rating tinggi dikategorikan sebagai sentimen positif, sedangkan ulasan dengan nilai rating rendah dikategorikan sebagai sentimen negatif.

Tabel 2. Data Positif (5 Baris Teratas) x

No	User Name	Score	Content
1	Pengguna google	1	Mohon maaf kenapa seiap kali ingin melihat ha..
2	Pengguna google	5	ulala

No	User Name	Score	Content
3	Pengguna google	5	Mantul
4	Pengguna google	5	Bagus
5	Pengguna google	3	Ini knp vidio shoopenya BLANK HITAM?

Tabel 3. Data Negatif (5 Baris Teratas)

No	UserName	Score	Content	Pelabelan
1	Pengguna goole	1	mohon maaf kenapa setiap kali ingin melihat ha...	negatif
2	Pengguna goole	1	Pembaruan aplikasi yang sangat buruk, tidak b...	negatif
3	Pengguna goole	1	Iklan k*****	negatif
4	Pengguna goole	2	Iklan shopee di web/aplikasi lain ngeselin	negatif
5	Pengguna goole	1	dapat notif paylater setelah diklik disuruh ng...	negatif

Tabel 2 dan Tabel 3 menunjukkan hasil pelabelan data ulasan produk Shopee yang dibagi ke dalam dua kategori sentimen, yaitu sentimen positif dan sentimen negatif. Pada bagian atas ditampilkan lima data teratas dengan sentimen positif, sedangkan pada bagian bawah ditampilkan lima data teratas dengan sentimen negatif. Setiap data ulasan memiliki beberapa atribut, yaitu userName, score, content, serta pelabelan sentimen.

Pelabelan sentimen dilakukan berdasarkan nilai skor dan isi ulasan yang diberikan oleh pengguna. Ulasan dengan penilaian tinggi dan isi yang bersifat mendukung dikategorikan sebagai sentimen positif, sedangkan ulasan dengan penilaian rendah dan isi yang mengandung keluhan atau kritik dikategorikan sebagai sentimen negatif. Proses pelabelan ini bertujuan untuk memudahkan analisis dan visualisasi data sentimen pelanggan pada tahap selanjutnya.

3.3 Text Pre-Processing

Proses text pre-processing adalah langkah penting dalam mempersiapkan data sebelum digunakan dalam sebuah model, yang bertujuan untuk mempersiapkan dokumen teks agar dapat diolah pada langkah selanjutnya.

Tabel 4. Data Positif (5 Baris Teratas dengan Tokenisasi)

No	UserName	Score	Content	Clean	Tokenized
1	Pengguna google	5	Ulala	Ulala	[ulala]
2	Pengguna google	5	Mantul	Mantul	[mantul]
3	Pengguna google	5	Bagus	Bagus	[bagus]
4	Pengguna google	5	good	Good	[Good]
5	Pengguna google	5	Sudah di-update malah close sendiri	sudah diupdet malah close sendiri	[Sudah, diUpdets, malah, close, sendiri]

Tabel 5. Data Negatif (5 Baris Teratas dengan Tokenisasi)

No	UserName	Score	contect	Clean text	tokenized
1	Pengguna google	1	mohon maaf kenapa setiap	mohon maaf kenapa setiap kali ingin melihat ha...	[mohon, maaf, kenapa, setiap,

No	UserName	Score	content	Clean text	tokenized
			kali ingin melihat ha...		kali, ingin, melihat, ha...
2	Pengguna google	1	Pembaruan aplikasi yang sangat buruk, tidak b...	pembaruan aplikasi yang sangat buruk tidak b...	[Pembaruan, aplikasi, yang, sangat, buruk, tidak, b...]
3	Pengguna google	1	iklan k*****	iklan k*****	[Iklan, k*****]
4	Pengguna google	2	Iklan shopee di web/aplikasi lain ngeselin	iklan shopee di web/aplikasi lain ngeselin	[Iklan, shopee, di, web/aplikasi, lain, ngeselin]
5	Pengguna google	1	dapat notif paylater setelah diklik disuruh n...	dapat notif paylater setelah diklik disuruh n...	[dapat, notif, paylater, setelah, diklik, disuruh, n...]

Tabel 4 dan Tabel 5 menampilkan hasil tokenisasi pada data ulasan produk Shopee yang telah melalui proses pembersihan teks (clean text). Data ditampilkan dalam dua kelompok, yaitu lima data teratas dengan sentimen positif dan lima data teratas dengan sentimen negatif. Setiap ulasan memiliki beberapa atribut, antara lain userName, score, content, pelabelan sentimen, clean text, dan tokenized. Pada tahap ini, teks ulasan yang masih mentah dibersihkan terlebih dahulu dengan menghilangkan karakter yang tidak diperlukan, kemudian dilakukan proses tokenisasi, yaitu pemecahan teks menjadi kata-kata tunggal. Hasil tokenisasi ditampilkan dalam bentuk daftar kata (list of tokens). Proses ini bertujuan untuk mempersiapkan data teks agar dapat digunakan pada tahap analisis dan visualisasi data selanjutnya.

3.3.1 Case folding

Case folding dilakukan untuk menghindari perbedaan makna akibat penggunaan huruf besar dan huruf kecil, sehingga kata yang sama dianggap identik dalam proses analisis teks.

Tabel 6. Tabel Case Folding

No	Content	Case Folding
1	mohon maaf kenapa setiap kali ingin melihat ha...	mohon maaf kenapa setiap kali ingin melihat ha...
2	Ulala	Ulala
3	mantul	mantul
4	Bagus	Bagus
5	Ini knp video shopeenya BLANK HITAM ?	ini knp video shopeenya blank hitam ?

Tabel 7 menunjukkan hasil proses case folding pada data ulasan produk Shopee. Pada tahap ini, teks ulasan asli (content) diubah menjadi huruf kecil seluruhnya pada kolom case_folding. Proses ini bertujuan untuk menyeragamkan format teks sehingga perbedaan penggunaan huruf besar dan huruf kecil tidak memengaruhi hasil analisis. Sebagai contoh, kata yang sebelumnya ditulis dengan huruf kapital seperti “BLANK HITAM” diubah menjadi “blank hitam”. Dengan adanya proses case folding, data teks menjadi lebih konsisten dan memudahkan proses selanjutnya seperti pembersihan teks, tokenisasi, serta analisis dan visualisasi data.

3.3.2 stop removal

Tabel 7. Data Positif (5 Baris Teratas setelah Stopword Removal)

No	Username	Score	Content	Pelabelan	Clean	Tokenized	Filtered Tokens
1	Pengguna google	5	Ulala	Positif	Ulala	[ulala]	[ulala]
2	Pengguna google	5	Mantul	Positif	Mantul	[mantul]	[mantul]
3	Pengguna google	5	Bagus	Positif	Bagus	[bagus]	[bagus]
4	Pengguna google	5	Good	Positif	Good	[Good]	[Good]
5	Pengguna google	5	Sudah diUpdets malah close sendiri	Positif	Sudah diUpdets malah close sendiri	[Sudah, diUpdets, malah, close, sendiri]	[Sudah, diUpdets, malah, close, sendiri]

Tabel 8. Data Negatif (3 Baris Teratas setelah Stopword Removal)

No	Usrnmme	Score	Content	Pelabelan	Clean	Tokenized	Filtere d_Tok ens
1	Pengguna google	1	mohon maaf kenapa setiap kali ingin melihat ha...	Negatif	mohon maaf kenapa setiap kali ingin melihat ha...	[mohon, maaf, kenapa, setiap, kali, ingin, mel...]	[mohon, maaf, halaman, beranda, jaringan...]
2	Pengguna google	1	Pembaruan aplikasi yang sangat buruk, tidak b...	Negatif	pembaruan aplikasi yang sangat buruk tidak b...	[Pembaruan, aplikasi, yang, sangat, buruk, ...]	[Pembaruan, aplikasi, buruk, pilih, ekspektasi...]
3	Pengguna google	1	Iklan k****	Negatif	Iklan k****	(iklan,kontoll)	(iklan, k****)

Tabel 8 dan tabel 9 menunjukkan hasil pelabelan sentimen serta proses text preprocessing pada data ulasan Shopee setelah tahap stopwords removal. Pada gambar tersebut ditampilkan contoh data dengan sentimen positif dan negatif berdasarkan nilai rating yang diberikan oleh pengguna. Kolom pelabelan_sentimen menunjukkan kategori sentimen dari setiap ulasan, sedangkan kolom clean text merupakan hasil pembersihan teks dari karakter tidak relevan. Selanjutnya, kolom tokenized menampilkan hasil pemecahan teks menjadi token kata, dan kolom filtered tokens menunjukkan token akhir yang telah melalui proses penyaringan kata tidak bermakna.

3.4 Ekstraksi Fitur

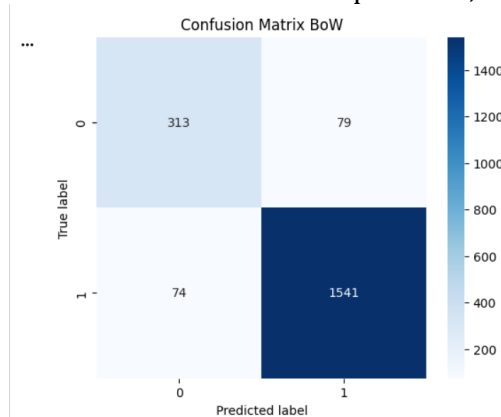
3.4.1 Bag Of Words (Bow)

Bag of Words merupakan metode ekstraksi fitur yang merepresentasikan teks sebagai kumpulan kata tanpa memperhatikan urutan kata. Setiap dokumen diubah menjadi vektor berdasarkan frekuensi kemunculan kata dalam dokumen tersebut, sehingga memudahkan proses analisis dan klasifikasi sentimen.

Tabel 9. Tabel BoW

Kelas / Rata-rata	Precision	Recall	F1-Score	Support
Negatif	0.81	0.88	0.80	392
Positif	0.95	0.95	0.95	1615
Accuracy	-	-	0.92	2007
Macro Avg	0.88	0.88	0.88	2007
Weighted Avg	0.92	0.92	0.92	2007

Hasil pengujian menunjukkan bahwa metode Bag of Words (BoW) menghasilkan akurasi sebesar 92,38%, dengan performa sangat baik pada kelas positif dan cukup seimbang pada kelas sentimen berdasarkan nilai precision, recall, dan F1-score.



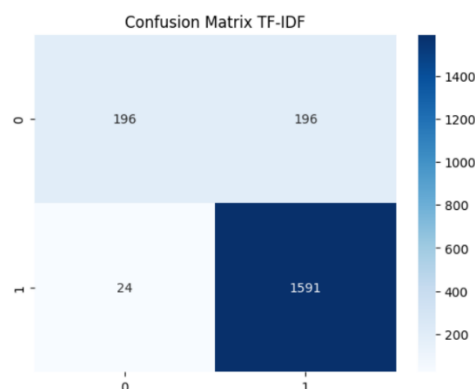
Gambar 2. Hasil Confusion Matrix

3.4.2 TF-IDF

Hasil pengujian menggunakan metode TF-IDF menunjukkan bahwa model mampu mengklasifikasikan sentiment dengan baik, ditunjukkan oleh nilai akurasi yang tinggi serta keseimbangan nilai precision, recall, dan F1-score pada masing-masing kelas sentiment.

Tabel 10. Tabel TF-IDF

Kelas / Rata-rata	Precision	Recall	F1-Score	Support
Negatif	0.89	0.50	0.64	392
Positif	0.89	0.99	0.94	1615
Accuracy	-	-	0.89	2007
Macro Avg	0.89	0.74	0.79	2007



Gambar 3. Hasil Confusion Matrix

```
print("Accuracy BoW :", accuracy_score(y_test, y_pred_bow))
print("Accuracy TF-IDF :", accuracy_score(y_test, y_pred_tfidf))
```

... Accuracy BoW : 0.9237668161434978
Accuracy TF-IDF : 0.8903836571998007

Gambar 4. Perbandingan hasil BoW dan TF-IDF

3.5 Klasifikasi Sentimen

3.5.1 SVM

Tabel 11. Hasil Evaluasi Model SVM

Kelas / Rata-rata	Precision	Recall	Fi-score	Support
Negatif	0.84	0.76	0.79	392
Positif	0.94	0.96	0.95	1615
Accuracy	-	-	0.92	2007
Macro Avg	0.89	0.86	0.87	2007
Weighted Avg	0.92	0.92	0.92	2007

Accuracy SVM : 0.923

Model Support Vector Machine (SVM) memperoleh nilai akurasi sebesar 92%, yang menunjukkan performa klasifikasi yang sangat baik. Pada sentimen positif, model menunjukkan kinerja tinggi dengan F1-score 0,95, sedangkan pada sentimen negatif diperoleh F1-score 0,79, yang menandakan peningkatan kemampuan deteksi dibandingkan model sebelumnya. Nilai weighted average F1-score sebesar 0,92 menunjukkan bahwa secara keseluruhan model SVM memiliki kinerja yang stabil dan seimbang dalam mengklasifikasikan sentimen ulasan.

3.5.2 Random Forest

Tabel 12. Hasil evaluasi model random forest

Kelas / Rata-rata	Precision	Recall	F1-Score	Support
Negatif	0.86	0.70	0.77	392
Positif	0.93	0.97	0.95	1615
Accuracy	—	—	0.92	2007
Macro Avg	0.89	0.83	0.86	2007
Weighted Avg	0.92	0.92	0.91	2007

Accuracy random Forest : 0.918

Model Random Forest menghasilkan nilai akurasi sebesar 92%, yang menunjukkan performa klasifikasi yang sangat baik. Pada sentimen positif, model mencapai F1-score 0,95, sedangkan pada sentimen negatif diperoleh F1-score 0,77, yang menunjukkan bahwa masih terdapat beberapa ulasan negatif yang belum terdeteksi secara optimal. Nilai weighted average F1-score sebesar 0,91 menandakan bahwa secara keseluruhan model Random Forest memiliki kinerja yang baik dan stabil dalam mengklasifikasikan sentimen ulasan.

3.5.3 Bernoulli naive bayes

Tabel 13. Model naive bayes

Kelas / Rata-rata	Precision	Recall	F1-Score	Support
Negatif	0.74	0.55	0.63	392
Positif	0.90	0.95	0.92	1615
Accuracy	-	-	0.87	2007
Macro Avg	0.82	0.75	0.78	2007
Weighted AVG	0.87	0.80.877	0.87	2007

Model Bernoulli Naive Bayes (BoW) menghasilkan akurasi 87%, artinya sebagian besar data berhasil diklasifikasikan dengan benar.

3.5.4 Perbandingan Hasil

Tabel 14. Perbandingan hasil kinerja Bow dan TF-IDF

Model	Accuracy	Precision	Recall	F1-Score
BoW	0.924	0.8880	0.876	0.878
TF-IDF	0.890	0.891	0.743	0.788

Tabel 15. Classification Report – BoW

Kelas / Rata-rata	Precision	Recall	F1-Score	Support
Negatif	0.81	0.80	0.80	392
Positif	0.95	0.95	0.95	1615
Accuracy	-	-	0.92	2007
Macro Avg	0.88	0.88	0.88	2007
Weighted Avg	0.92	0.92	0.92	2007

Tabel 16. Classification Report – TF-IDF

Kelas / Rata-rata	Precision	Recall	F1-Score	Support
Negatif	0.89	0.50	0.64	392
Positif	0.89	0.99	0.94	1615
Accuracy	-	-	0.89	2007
Macro Avg	0.89	0.74	0.79	2007
Weighted Avg	0.89	0.89	0.88	2007

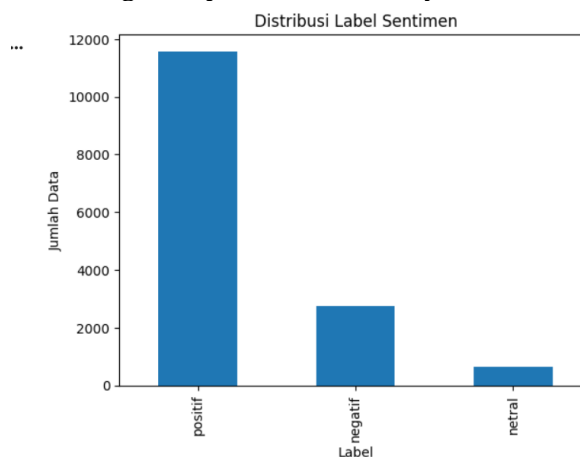
Perbandingan akurasi beberapa algoritma klasifikasi teks menggunakan fitur Bag of Words (bow) dan TF-IDF. Pada bow, Naive Bayes memberikan akurasi tertinggi ($\pm 92\%$), sedangkan pada TF-IDF, SVM menghasilkan akurasi terbaik ($\pm 93\%$). Secara umum, TF-IDF cenderung memberikan performa yang lebih stabil dan tinggi dibandingkan bow pada sebagian besar model.

3.6 Visualisasi

3.6.1 Visualisasi Distribusi Label

Gambar tersebut menampilkan distribusi label sentimen pada dataset ulasan Shopee yang terdiri dari sentimen positif negatif dan netral Sumbu horizontal menunjukkan kategori sentimen sedangkan sumbu vertikal menunjukkan jumlah data ulasan Berdasarkan grafik terlihat bahwa sentimen positif memiliki jumlah data paling besar dibandingkan dengan sentimen lainnya Kondisi ini menunjukkan bahwa sebagian besar pengguna Shopee memberikan ulasan yang bersifat positif terhadap produk maupun layanan yang diterima. Sentimen negatif memiliki jumlah data yang lebih sedikit dibandingkan sentimen positif namun masih cukup signifikan Ulasan negatif umumnya mencerminkan ketidakpuasan pengguna yang berkaitan dengan kualitas produk keterlambatan pengiriman atau pelayanan penjual.

Sementara itu sentimen netral merupakan kategori dengan jumlah data paling sedikit Ulasan netral biasanya berisi pernyataan singkat atau informasi tanpa menunjukkan kecenderungan emosi yang jelas baik positif maupun negatif. Distribusi data pada grafik tersebut menunjukkan adanya ketidakseimbangan kelas di mana data sentimen positif jauh lebih dominan dibandingkan sentimen negatif dan netral Ketidakseimbangan ini perlu diperhatikan dalam proses analisis sentimen karena dapat memengaruhi kinerja model klasifikasi yang cenderung memprediksi kelas mayoritas.



Gambar 5. Label sentimen distribusi

4. Kesimpulan

Hasil pengujian menunjukkan bahwa metode Bag of Words (BoW) menghasilkan performa yang lebih baik dibandingkan TF-IDF pada dataset yang digunakan, dengan tingkat akurasi mencapai sekitar 92%. Di antara algoritma klasifikasi yang diterapkan, Support Vector Machine (SVM) dan Random Forest menunjukkan kinerja yang paling baik dan stabil dalam mengklasifikasikan sentimen positif dan negatif. Sementara itu, algoritma Bernoulli Naive Bayes juga mampu melakukan klasifikasi dengan cukup baik meskipun menghasilkan akurasi yang lebih rendah dibandingkan algoritma lainnya.

Selain itu, hasil visualisasi data menunjukkan bahwa mayoritas ulasan pengguna Shopee bersifat positif, yang mengindikasikan tingkat kepuasan pengguna yang cukup tinggi terhadap produk dan layanan yang disediakan. Namun, ditemukan pula ketidakseimbangan kelas pada dataset, di mana jumlah ulasan positif jauh lebih dominan dibandingkan ulasan negatif. Secara keseluruhan, penelitian ini membuktikan bahwa pendekatan text mining dan klasifikasi teks dapat digunakan sebagai alat bantu yang efektif untuk memahami persepsi dan kepuasan pelanggan pada platform e-commerce

Daftar Pustaka

- [1] M. S. Suib and V. Y. Habsari, "Persepsi Customer terhadap Platform E-Commerce Shopee," *Jurnal Ekonomi dan Pendidikan*, vol. 19, no. 2, pp. 1–10, 2022, doi: <https://doi.org/10.21831/jep.v19i2.48023>
 - [2] A. P. Nellutla, M. Hudnurkar, S. S. Ambekar, and A. D. Lidbe, "Online product reviews and their impact on third party sellers using natural language processing," *International Journal of Business Intelligence Research*, vol. 12, no. 1, pp. 26–47, Jan. 2021, doi: <https://doi.org/10.4018/IJBIR.20210101.0a2>.
 - [3] M. Harsanto and E. Sudarmilah, "TINJAUAN LITERATUR ANALISIS SENTIMEN PRODUK E-COMMERCE: DATASET, PENDEKATAN, METODE, DAN PERFORMA," *JUPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 10, no. 3, pp. 2290–2303, Aug. 2025, doi: <https://doi.org/10.29100/jupi.v10i3.8217>.
 - [4] M. D. Rizkiyanto, M. D. Purbolaksono, and W. Astuti, "Sentiment Analysis Classification on PLN Mobile Application Reviews using Random Forest Method and TF-IDF Feature Extraction," *INTEK: Jurnal Penelitian*, vol. 11, no. 1, pp. 37–43, Apr. 2024, doi: <https://doi.org/10.31963/intek.v11i1.4774>
 - [5] R. C. Rivaldi and T. D. Wismarini, "Analisis Sentimen Pada Ulasan Produk Dengan Metode Natural Language Processing (NLP)," *Elkom: Jurnal Elektronika dan Komputer*, vol. 17, no. 1, pp. 120–128, Jul. 2024, doi: <https://doi.org/10.51903/elkom.v17i1.1680>
 - [6] S.Muthukumaran, Dr.P.Suresh, "Text Analysis for Product Reviews for Sentiment Analysis using NLP Methods," *International Journal of Engineering Trends and Technology (IJETT)*, vol. 47, no. 4, pp. 474–480, 2017. Crossref, <https://doi.org/10.14445/22315381/IJETT-V47P278>
 - [7] E. Halim and R. Purba, "Consumer Opinion Extraction Using Text Mining for Product Recommendations On E-Commerce," *Indonesian Journal of Artificial Intelligence and Data Mining (IJAIMD)*, vol. 4, no. 1, pp. 19–28, 2021, doi: <https://doi.org/10.24014/ijaidm.v4i1.10834>
 - [8] A. Alaiya and C. Agusniar, "Sentiment Analysis of E-Commerce Product Reviews on Tokopedia Using Support Vector Machine," 2025. [Online]. Available: <https://doi.org/10.30871/jaic.v9i5.10977>
 - [9] N. Sikana, S. Winardi, G. -, G. F. Situmorang, and R. Lubis, "Analisis Sentimen untuk Ulasan Produk E-Commerce Shopee Menggunakan BERT," *Jurnal Sifo Mikroskil*, vol. 26, no. 2, Oct. 2025, doi: <https://doi.org/10.55601/jsm.v26i2.1796>
 - [10] J. Angreyani and Y. Pernando, "ANALISIS SENTIMEN ULASAN E-COMMERCE SHOPEE DENGAN MENGGUNAKAN ALGORITMA NAIVE BAYES," *J-Com (Journal of Computer)*, vol. 5, no. 1, pp. 1–8, Mar. 2025, doi: <https://doi.org/10.59934/jaiea.v4i3.1132>
-