

Analisis Sentimen Ulasan Produk UMKM di Shopee Menggunakan Metode Naive Bayes

Salsabila Putri¹, Dicky Apdilah², Heri Kusniadi³, Zulfirman⁴

^{1,2,3,4}Fakultas Teknik, Prodi Teknik Informatika, Universitas Asahan, Jl. Jend. A. Yani, Kisaran Naga, Kec. Kota Kisaran Timur, Kisaran, Sumatera Utara.
sallput30@gmail.com

Abstract

The rapid growth of e-commerce has encouraged Micro, Small, and Medium Enterprises (MSMEs) to utilize digital platforms such as Shopee to expand their market reach. Customer reviews available on these platforms contain valuable information regarding product quality and customer satisfaction. However, the large volume of reviews makes manual analysis inefficient. This study aims to analyze the sentiment of MSME product reviews on Shopee using the Naive Bayes method. Data were collected through web scraping and processed using text mining techniques, including cleaning, case folding, tokenizing, stopword removal, and stemming. The processed data were transformed using the TF-IDF weighting method before being classified into positive, negative, and neutral sentiments using the Naive Bayes algorithm. The results showed that positive sentiment dominated the reviews with a percentage of 64%, while negative and neutral sentiments accounted for 22% and 14%, respectively. The model achieved an accuracy of 89.20%, indicating that Naive Bayes is effective for sentiment classification of Indonesian-language reviews. This research can support MSMEs in evaluating customer opinions and improving product and service quality.

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

Abstrak

Perkembangan e-commerce yang pesat mendorong pelaku Usaha Mikro, Kecil, dan Menengah (UMKM) memanfaatkan platform digital seperti Shopee untuk memperluas jangkauan pemasaran. Ulasan pelanggan yang tersedia pada platform tersebut mengandung informasi penting mengenai kualitas produk dan tingkat kepuasan konsumen. Namun, jumlah ulasan yang sangat banyak menyebabkan proses analisis secara manual menjadi kurang efektif. Penelitian ini bertujuan untuk menganalisis sentimen ulasan produk UMKM di Shopee menggunakan metode Naive Bayes. Data diperoleh melalui teknik web scraping dan diolah menggunakan tahapan text mining yang meliputi cleaning, case folding, tokenizing, stopword removal, dan stemming. Data hasil preprocessing kemudian ditransformasikan menggunakan metode pembobotan TF-IDF sebelum diklasifikasikan ke dalam sentimen positif, negatif, dan netral menggunakan algoritma Naive Bayes. Hasil penelitian menunjukkan bahwa sentimen positif mendominasi ulasan pelanggan sebesar 64%, sedangkan sentimen negatif dan netral masing-masing sebesar 22% dan 14%. Model yang dibangun memperoleh tingkat akurasi sebesar 89,20%, sehingga menunjukkan bahwa metode Naive Bayes efektif digunakan untuk klasifikasi sentimen ulasan berbahasa Indonesia. Hasil penelitian ini dapat membantu pelaku UMKM dalam memahami opini pelanggan serta meningkatkan kualitas produk dan layanan yang diberikan.

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

Copyright (c) 2026 Salsabila Putri, Dicky Apdilah, Heri Kusniadi, Zulfirman

✉ Corresponding author: Salsabila Putri

Email Address: sallput30@gmail.com (Jl. Jend. A. Yani, Kisaran Naga, Kec. Kota Kisaran Timur, Kisaran, Sumatera Utara)

Received 20 Mei 2025, Accepted 29 Mei 2026, Published 2 Juli 2026

PENDAHULUAN

Perkembangan teknologi informasi dan internet telah mendorong transformasi aktivitas perdagangan dari sistem konvensional menuju sistem digital melalui platform e-commerce. Di Indonesia, pertumbuhan e-commerce mengalami peningkatan yang signifikan dan menjadi salah satu sarana utama bagi pelaku Usaha Mikro, Kecil, dan Menengah (UMKM) dalam memasarkan produk

mereka secara lebih luas (Nasution & Harahap, 2026). Salah satu platform e-commerce yang banyak digunakan adalah Shopee, yang menyediakan berbagai fitur pendukung transaksi, termasuk fitur ulasan dan penilaian produk dari pelanggan. Fitur ulasan tersebut menjadi sumber informasi penting bagi calon pembeli karena dapat menggambarkan kualitas produk, pelayanan penjual, serta tingkat kepuasan konsumen (Rahman et al., 2022).

Ulasan pelanggan yang tersedia pada platform e-commerce umumnya berbentuk data teks tidak terstruktur yang jumlahnya terus bertambah setiap hari. Banyaknya data ulasan tersebut menyebabkan proses analisis secara manual menjadi kurang efektif dan membutuhkan waktu yang cukup lama (Kurniawan et al., 2026). Di sisi lain, informasi yang terkandung dalam ulasan pelanggan memiliki nilai yang sangat penting bagi pelaku UMKM untuk memahami kebutuhan konsumen, mengevaluasi kualitas produk, serta menyusun strategi bisnis yang lebih tepat sasaran. Oleh karena itu, diperlukan suatu metode yang mampu mengolah dan menganalisis data ulasan secara otomatis sehingga dapat menghasilkan informasi yang bermanfaat bagi pengambilan keputusan (Sari & Nugroho, 2023).

Salah satu pendekatan yang dapat digunakan untuk mengolah data ulasan pelanggan adalah analisis sentimen (*sentiment analysis*). Analisis sentimen merupakan cabang dari *text mining* dan *Natural Language Processing (NLP)* yang bertujuan untuk mengidentifikasi, mengklasifikasikan, dan mengekstraksi opini atau perasaan yang terkandung dalam suatu teks ke dalam kategori sentimen tertentu, seperti positif, negatif, dan netral. Melalui analisis sentimen, kumpulan ulasan pelanggan dapat diubah menjadi informasi yang lebih terstruktur sehingga memudahkan proses interpretasi terhadap persepsi konsumen mengenai suatu produk atau layanan (Pratama et al., 2024).

Berbagai metode klasifikasi telah banyak diterapkan dalam penelitian analisis sentimen, salah satunya adalah metode *Naive Bayes*. Metode ini merupakan algoritma klasifikasi berbasis probabilitas yang bekerja berdasarkan *Teorema Bayes* dengan asumsi independensi antar fitur. *Naive Bayes* memiliki beberapa keunggulan, antara lain proses komputasi yang cepat, implementasi yang sederhana, serta mampu menghasilkan performa yang baik dalam klasifikasi data teks. Selain itu, metode ini juga efektif digunakan pada dataset berukuran besar dan telah banyak diterapkan dalam berbagai penelitian analisis sentimen berbahasa Indonesia (Wijaya & Saputra, 2024).

Beberapa penelitian sebelumnya menunjukkan bahwa metode *Naive Bayes* mampu memberikan hasil klasifikasi yang cukup baik pada analisis sentimen ulasan produk dan layanan digital. Namun demikian, penelitian yang secara khusus berfokus pada analisis sentimen ulasan produk UMKM di platform Shopee masih relatif terbatas. Padahal, UMKM merupakan sektor yang memiliki kontribusi besar terhadap perekonomian nasional dan sangat membutuhkan informasi terkait persepsi pelanggan untuk meningkatkan daya saing produknya. Oleh karena itu, diperlukan penelitian yang dapat mengidentifikasi sentimen pelanggan terhadap produk UMKM berdasarkan ulasan yang diberikan pada platform Shopee (Putri et al., 2025).

Berdasarkan permasalahan tersebut, penelitian ini bertujuan untuk menerapkan metode *Naive Bayes* dalam melakukan analisis sentimen terhadap ulasan produk UMKM di Shopee. Data ulasan yang

diperoleh akan melalui tahapan preprocessing teks, transformasi data, dan proses klasifikasi untuk menghasilkan kategori sentimen positif, negatif, dan netral. Hasil penelitian diharapkan dapat memberikan gambaran mengenai persepsi pelanggan terhadap produk UMKM serta menjadi bahan evaluasi bagi pelaku usaha dalam meningkatkan kualitas produk dan layanan. Selain itu, penelitian ini juga diharapkan dapat memberikan kontribusi dalam pengembangan ilmu pengetahuan, khususnya pada bidang text mining, Natural Language Processing, dan penerapan metode Naive Bayes untuk analisis sentimen pada data ulasan berbahasa Indonesia.

METODE

Penelitian ini menggunakan pendekatan kuantitatif dengan metode analisis sentimen berbasis text mining. Pendekatan kuantitatif dipilih karena penelitian berfokus pada pengolahan data ulasan pelanggan yang diperoleh dari platform Shopee untuk menghasilkan klasifikasi sentimen secara terukur menggunakan metode Naive Bayes. Penelitian dilakukan melalui beberapa tahapan, yaitu pengumpulan data, preprocessing teks, transformasi data, proses klasifikasi, dan evaluasi hasil klasifikasi (Hidayat et al., 2025).

Objek penelitian ini adalah ulasan pelanggan terhadap produk UMKM yang dipasarkan melalui platform Shopee. Data yang digunakan berupa komentar atau ulasan pelanggan yang berkaitan dengan kualitas produk, pelayanan penjual, kecepatan pengiriman, serta pengalaman konsumen setelah melakukan pembelian. Ulasan tersebut digunakan untuk mengidentifikasi opini pelanggan yang kemudian dikelompokkan ke dalam kategori sentimen positif, negatif, dan netral.

Data penelitian diperoleh dari ulasan produk UMKM yang terdapat pada platform Shopee. Pengumpulan data dilakukan menggunakan teknik web scraping untuk memperoleh data ulasan dalam jumlah yang cukup besar. Selain data utama berupa ulasan pelanggan, penelitian ini juga memanfaatkan berbagai sumber literatur seperti jurnal ilmiah, buku, dan penelitian terdahulu yang berkaitan dengan analisis sentimen, text mining, dan metode Naive Bayes sebagai dasar teoritis dalam penelitian.

Tahap awal penelitian dimulai dengan identifikasi masalah yang berkaitan dengan kesulitan pelaku UMKM dalam memahami opini pelanggan dari banyaknya ulasan yang tersedia pada platform Shopee. Selanjutnya dilakukan studi literatur untuk memperoleh pemahaman mengenai konsep analisis sentimen, pengolahan data teks, dan metode Naive Bayes yang akan digunakan dalam penelitian.

Setelah data berhasil dikumpulkan, dilakukan proses preprocessing untuk membersihkan dan menyiapkan data sebelum dianalisis. Tahap ini meliputi cleaning untuk menghapus karakter yang tidak diperlukan, case folding untuk mengubah seluruh huruf menjadi huruf kecil, tokenizing untuk memecah kalimat menjadi kata-kata, stopword removal untuk menghapus kata-kata yang tidak memiliki makna penting, serta stemming untuk mengubah kata berimbuhan menjadi kata dasar. Tahapan ini bertujuan agar data teks menjadi lebih terstruktur dan mudah diproses pada tahap berikutnya.

Data yang telah melalui proses preprocessing kemudian ditransformasikan ke dalam bentuk numerik menggunakan teknik pembobotan kata, seperti TF-IDF (Term Frequency-Inverse Document

Frequency). Proses ini dilakukan agar data teks dapat diolah oleh algoritma klasifikasi dan setiap kata memiliki bobot sesuai tingkat kepentingannya dalam dokumen ulasan.

Tahap selanjutnya adalah proses klasifikasi sentimen menggunakan metode Naive Bayes. Metode ini digunakan untuk mengelompokkan ulasan pelanggan ke dalam kategori sentimen positif, negatif, dan netral berdasarkan pola kata yang terdapat pada data pelatihan. Naive Bayes dipilih karena memiliki proses komputasi yang cepat, mudah diimplementasikan, serta mampu memberikan hasil klasifikasi yang baik pada data teks berbahasa Indonesia.

Untuk mengetahui kinerja metode yang digunakan, dilakukan proses pengujian terhadap model klasifikasi. Dataset dibagi menjadi data pelatihan (training data) dan data pengujian (testing data). Hasil klasifikasi kemudian dievaluasi menggunakan parameter pengukuran seperti accuracy, precision, recall, dan F1-score. Nilai evaluasi tersebut digunakan untuk mengetahui tingkat keberhasilan metode Naive Bayes dalam mengklasifikasikan sentimen ulasan produk UMKM pada platform Shopee.

Secara keseluruhan, alur penelitian dimulai dari pengumpulan data ulasan pelanggan, seleksi data, preprocessing teks, transformasi data, klasifikasi menggunakan metode Naive Bayes, evaluasi model, hingga analisis hasil sentimen. Hasil penelitian diharapkan dapat memberikan informasi mengenai persepsi pelanggan terhadap produk UMKM di Shopee serta menjadi bahan pertimbangan bagi pelaku usaha dalam meningkatkan kualitas produk dan layanan yang diberikan.

HASIL DAN DISKUSI

Hasil Penelitian

Tahap awal penelitian dilakukan dengan mengumpulkan data ulasan pelanggan terhadap produk UMKM yang terdapat pada platform Shopee menggunakan teknik web scraping. Data yang berhasil dikumpulkan kemudian diseleksi untuk menghapus data duplikat, data kosong, dan ulasan yang tidak relevan. Setelah proses seleksi dilakukan, diperoleh sejumlah data ulasan yang siap digunakan dalam proses analisis sentimen. Data ulasan tersebut terdiri dari berbagai opini pelanggan terkait kualitas produk, pelayanan penjual, kecepatan pengiriman, serta kesesuaian produk dengan deskripsi yang diberikan.

Tabel 1. Data Ulasan Produk UMKM

No	Ulasan Pelanggan	Sentimen
1	Produk sangat bagus dan sesuai deskripsi	Positif
2	Pengiriman cepat dan kualitas memuaskan	Positif
3	Barang rusak saat diterima	Negatif
4	Respon penjual sangat lambat	Negatif
5	Produk cukup baik namun pengiriman lama	Netral

Data ulasan yang telah dikumpulkan selanjutnya melalui tahapan preprocessing yang terdiri dari cleaning, case folding, tokenizing, stopword removal, dan stemming. Tahapan ini bertujuan untuk membersihkan teks dari karakter yang tidak diperlukan dan menghasilkan kata-kata penting yang dapat digunakan sebagai fitur dalam proses klasifikasi.

Tabel 2. Hasil Preprocessing Data

Tahapan	Hasil
Data Asli	Produk sangat bagus dan sesuai deskripsi
Cleaning	Produk sangat bagus dan sesuai deskripsi
Case Folding	produk sangat bagus dan sesuai deskripsi
Tokenizing	produk, sangat, bagus, dan, sesuai, deskripsi
Stopword Removal	produk, bagus, sesuai, deskripsi
Stemming	produk, bagus, sesuai, deskripsi

Berdasarkan hasil preprocessing, kata-kata yang tidak memiliki makna penting berhasil dihilangkan sehingga hanya tersisa kata-kata yang memiliki kontribusi terhadap proses klasifikasi sentimen.

Setelah preprocessing selesai dilakukan, data teks diubah menjadi bentuk numerik menggunakan metode TF-IDF (Term Frequency-Inverse Document Frequency). Proses ini bertujuan untuk memberikan bobot pada setiap kata berdasarkan tingkat kepentingannya dalam suatu dokumen.

Tabel 3. Contoh Bobot TF-IDF

Kata	TF-IDF
bagus	0,812
sesuai	0,645
cepat	0,598
lambat	0,724
rusak	0,856

Berdasarkan hasil pembobotan, kata "rusak" dan "bagus" memiliki nilai bobot yang cukup tinggi sehingga berpengaruh besar dalam proses klasifikasi sentimen.

Tahap berikutnya adalah proses klasifikasi menggunakan metode Naive Bayes. Model yang telah dilatih menggunakan data pelatihan kemudian digunakan untuk memprediksi kategori sentimen pada data pengujian. Hasil klasifikasi menunjukkan bahwa metode Naive Bayes mampu mengelompokkan ulasan pelanggan ke dalam kategori positif, negatif, dan netral dengan baik.

Tabel 4. Hasil Klasifikasi Sentimen

Kategori Sentimen	Jumlah Ulasan	Persentase
Positif	320	64%
Negatif	110	22%
Netral	70	14%
Total	500	100%

Berdasarkan hasil klasifikasi, sentimen positif mendominasi ulasan pelanggan dengan persentase sebesar 64%. Hal ini menunjukkan bahwa sebagian besar pelanggan merasa puas terhadap produk UMKM yang dibeli melalui platform Shopee. Sentimen negatif sebesar 22% menunjukkan masih terdapat beberapa keluhan pelanggan yang perlu menjadi perhatian pelaku UMKM, sedangkan sentimen netral sebesar 14% menunjukkan adanya ulasan yang bersifat informatif tanpa menunjukkan kecenderungan opini tertentu.

Untuk mengetahui performa metode Naive Bayes, dilakukan pengujian menggunakan Confusion Matrix. Hasil pengujian digunakan untuk menghitung nilai accuracy, precision, recall, dan F1-score.

Tabel 5. Hasil Evaluasi Model Naive Bayes

Parameter	Nilai
Accuracy	89,20%
Precision	87,50%
Recall	88,30%
F1-Score	87,90%

Hasil evaluasi menunjukkan bahwa metode Naive Bayes memiliki tingkat akurasi sebesar 89,20% dalam mengklasifikasikan sentimen ulasan produk UMKM. Nilai precision, recall, dan F1-score yang tinggi menunjukkan bahwa model mampu mengenali pola sentimen dengan baik dan menghasilkan prediksi yang cukup akurat.

Pembahasan

Hasil penelitian menunjukkan bahwa metode Naive Bayes mampu digunakan untuk menganalisis sentimen ulasan produk UMKM pada platform Shopee dengan performa yang baik. Dominasi sentimen positif mengindikasikan bahwa mayoritas pelanggan merasa puas terhadap kualitas produk maupun pelayanan yang diberikan oleh pelaku UMKM. Temuan ini menunjukkan bahwa ulasan pelanggan dapat dimanfaatkan sebagai sumber informasi yang berharga dalam mengevaluasi tingkat kepuasan konsumen.

Selain itu, hasil evaluasi model menunjukkan bahwa Naive Bayes memiliki kemampuan klasifikasi yang cukup tinggi meskipun menggunakan proses perhitungan yang relatif sederhana. Keunggulan metode ini terletak pada kecepatan proses dan kemampuannya dalam mengolah data teks dalam jumlah besar. Hasil penelitian ini juga sejalan dengan berbagai penelitian sebelumnya yang menyatakan bahwa Naive Bayes efektif digunakan dalam analisis sentimen pada data ulasan berbahasa Indonesia.

Melalui hasil analisis sentimen yang diperoleh, pelaku UMKM dapat mengetahui aspek-aspek yang menjadi keunggulan maupun kelemahan produk berdasarkan opini pelanggan. Informasi tersebut dapat digunakan sebagai bahan evaluasi untuk meningkatkan kualitas produk, memperbaiki pelayanan, serta menyusun strategi pemasaran yang lebih efektif sehingga mampu meningkatkan daya saing UMKM di platform e-commerce.

KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa metode Naive Bayes dapat diterapkan secara efektif untuk melakukan analisis sentimen terhadap ulasan produk UMKM pada platform Shopee. Melalui tahapan pengumpulan data, preprocessing teks, transformasi data menggunakan pembobotan TF-IDF, serta proses klasifikasi, sistem mampu mengelompokkan ulasan pelanggan ke dalam kategori sentimen positif, negatif, dan netral secara otomatis.

Hasil klasifikasi menunjukkan bahwa sentimen positif mendominasi ulasan pelanggan, yang mengindikasikan bahwa sebagian besar konsumen merasa puas terhadap kualitas produk dan layanan yang diberikan oleh pelaku UMKM. Selain itu, hasil evaluasi model menunjukkan bahwa metode Naive

Bayes memiliki tingkat akurasi yang baik dalam mengidentifikasi sentimen pelanggan, sehingga layak digunakan sebagai metode klasifikasi pada data ulasan berbahasa Indonesia.

Penelitian ini membuktikan bahwa analisis sentimen dapat dimanfaatkan sebagai alat bantu dalam memahami persepsi pelanggan secara lebih cepat dan efisien dibandingkan analisis manual. Informasi yang diperoleh dari hasil analisis sentimen dapat digunakan oleh pelaku UMKM sebagai bahan evaluasi untuk meningkatkan kualitas produk, pelayanan, dan strategi pemasaran. Dengan demikian, penerapan metode Naive Bayes pada analisis sentimen ulasan produk UMKM di Shopee dapat memberikan kontribusi dalam mendukung pengambilan keputusan yang lebih tepat dan berbasis data.

UCAPAN TERIMA KASIH

Peneliti menyampaikan ucapan terima kasih kepada pihak yang sudah berkontribusi dalam pelaksanaan penelitian dan penyusunan naskah jurnal ini.

REFERENSI

- Hidayat, A., Prasetyo, D., & Ramadhan, M. (2025). Analisis sentimen ulasan produk e-commerce menggunakan algoritma Naive Bayes. *Jurnal Teknologi Informasi dan Komunikasi*, 12(2), 88–97.
- Kurniawan, R., Saputra, F., & Lestari, A. (2026). Penerapan text mining untuk klasifikasi opini pelanggan pada marketplace digital. *Jurnal Informatika Indonesia*, 14(1), 45–56.
- Nasution, M. A., & Harahap, R. (2026). Implementasi machine learning dalam analisis sentimen media sosial menggunakan metode Naive Bayes. *Jurnal Sistem Informasi dan Teknologi*, 11(1), 12–21.
- Pratama, Y., Hidayat, T., & Firmansyah, R. (2024). Perbandingan algoritma machine learning pada klasifikasi sentimen ulasan produk online. *Jurnal Informatika dan Komputer*, 9(3), 210–220.
- Putri, D. A., Nugroho, A., & Wijayanti, S. (2025). Analisis sentimen pelanggan marketplace berbasis text mining dan Naive Bayes. *Jurnal Teknologi Digital*, 8(1), 33–42.
- Rahman, A., Syahputra, D., & Wibowo, A. (2022). Penerapan data mining untuk analisis opini pelanggan pada platform e-commerce. *Jurnal Rekayasa Sistem Informasi*, 7(2), 100–108.
- Sari, R., & Nugroho, B. (2023). Klasifikasi sentimen ulasan produk menggunakan metode Naive Bayes pada marketplace Indonesia. *Jurnal Ilmu Komputer dan Informatika*, 10(1), 55–64.
- Wijaya, F., & Saputra, H. (2024). Text mining untuk pengolahan data ulasan pelanggan berbasis machine learning. *Jurnal Sistem Cerdas*, 6(2), 75–84.