

Analisis Sentimen Ulasan Pengguna Shopee Terkait Fitur *Cash on Delivery* Menggunakan Algoritma Naive Bayes

Ani Sofiaturo Rosidah¹, Aisyah Lia Sari¹, Fitria Apriani¹

¹Universitas Nurul Huda
E-mail: sofearoseda@gmail.com

Abstrak

Perkembangan *e-commerce* di Indonesia mendorong pemanfaatan fitur *Cash on Delivery* (COD) sebagai solusi inklusi finansial. Namun, dinamika fitur ini diwarnai berbagai konflik lapangan dan penolakan paket yang tercermin dalam ulasan pengguna di platform Shopee. Penelitian ini bertujuan untuk menganalisis sentimen pengguna Shopee terkait fitur COD menggunakan algoritma *Multinomial Naive Bayes* berbasis pembobotan *Term Frequency-Inverse Document Frequency* (TF-IDF). Sebanyak 1.000 ulasan sekunder dari Kaggle disaring secara biner menjadi 546 sentimen negatif dan 454 sentimen positif. Dataset dibagi menggunakan metode *Hold-Out* dengan proporsi rasio 80:20 menjadi 800 data latih dan 200 data uji. Hasil pengujian menunjukkan model mampu menghasilkan tingkat akurasi global sebesar 78,00%. Melalui analisis *confusion matrix*, diketahui bahwa algoritma memiliki sensitivitas yang sangat tinggi dalam mengenali keluhan pengguna dengan nilai *recall* sentimen negatif mencapai 90%, sedangkan *recall* sentimen positif sebesar 62%. Tingginya kemampuan deteksi teks negatif ini dipicu oleh karakteristik ulasan komplain masyarakat Indonesia yang cenderung ekspresif dan repetitif. Penelitian ini membuktikan bahwa kombinasi TF-IDF dan *Multinomial Naive Bayes* sangat layak diandalkan oleh manajemen *e-commerce* sebagai sistem pemantauan kualitas layanan otomatis.

Kata kunci : Analisis Sentimen, *Cash on Delivery*, *Multinomial Naive Bayes*, Shopee, TF-IDF.

Abstract

The development of e-commerce in Indonesia has encouraged the use of Cash on Delivery (COD) features as a financial inclusion solution. However, the dynamics of this feature are colored by various field conflicts and package rejections, which are reflected in user reviews on the Shopee platform. This study aims to analyze Shopee user sentiment regarding the COD feature using the Multinomial Naive Bayes algorithm based on Term Frequency-Inverse Document Frequency (TF-IDF) weighting. A total of 1,000 secondary reviews from Kaggle were binarily filtered into 546 negative sentiments and 454 positive sentiments. The dataset was split using the Hold-Out method with an 80:20 ratio into 800 training data and 200 testing data. The test results show that the model is capable of producing a global accuracy rate of 78.00%. Through confusion matrix analysis, it is known that the algorithm has a very high sensitivity in recognizing user complaints, with the recall value for negative sentiment reaching 90%, while the positive sentiment recall is 62%. The high capability of negative text detection is triggered by the characteristics of Indonesian complaint reviews, which tend to be expressive and repetitive. This study proves that the combination of TF-IDF and Multinomial Naive Bayes is highly reliable for e-commerce management as an automatic service quality monitoring system

Keywords: Sentiment Analysis, *Cash on Delivery*, *Multinomial Naive Bayes*, Shopee, TF-IDF.

PENDAHULUAN

Perkembangan teknologi informasi mendorong pertumbuhan *electronic commerce* (*e-commerce*) di Indonesia. Shopee, sebagai salah satu platform utama, menyediakan berbagai metode pembayaran untuk mempermudah transaksi, termasuk fitur *Cash on Delivery* (COD) [1]. Di Indonesia, COD memiliki karakteristik unik karena menjadi solusi inklusi finansial bagi masyarakat yang belum

memiliki akses perbankan (*unbanked*) serta mengatasi krisis kepercayaan dalam belanja daring. Namun, dinamika sistem COD saat ini diwarnai masalah spesifik yang kompleks, seperti fenomena viral penolakan paket sepihak, konflik fisik antara pembeli dan kurir akibat kesalahpahaman regulasi, hingga tingginya angka pengembalian barang (*return rate*). Dinamika emosional ini menghasilkan ulasan pengguna yang kaya akan variasi kosakata, repetisi kata makian atau pujian, serta sarkasme, yang sangat krusial untuk dievaluasi [2].

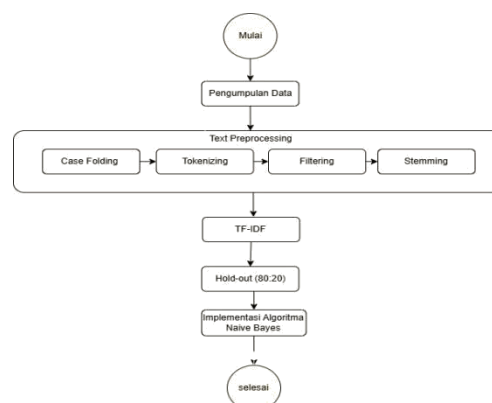
Meningkatnya jumlah ulasan menyebabkan analisis manual tidak lagi efektif. Oleh karena itu, diperlukan teknik analisis sentimen (*sentiment analysis*) berbasis *Natural Language Processing* (NLP) untuk mengklasifikasikan opini ke dalam sentimen positif dan negatif secara otomatis [3]. Beberapa penelitian terdahulu telah menerapkan algoritma Naive Bayes untuk ulasan *e-commerce* [4]. Namun, penelitian yang spesifik membahas dinamika fitur COD masih terbatas. Selain itu, terdapat *research gap* di mana studi terdahulu jarang memberikan landasan empiris dalam pemilihan varian Naive Bayes yang digunakan pada karakteristik data ulasan tertentu.

Karakteristik data ulasan COD pada platform Shopee secara spesifik berupa teks dokumen diskrit dengan panjang kalimat bervariasi yang mengandung kemunculan kata emosional berulang (misal: kata "kecewa" atau "kurir" muncul beberapa kali dalam satu ulasan). Berdasarkan data empiris tersebut, varian *Gaussian Naive Bayes* tidak relevan karena mensyaratkan data numerik kontinu yang berdistribusi normal. Varian *Bernoulli Naive Bayes* juga tidak tepat karena hanya mendeteksi keberadaan kata secara biner (0 atau 1), sehingga membuang informasi bobot frekuensi kata emosional yang ditulis berulang oleh pengguna. Oleh karena itu, penelitian ini menggunakan *Multinomial Naive Bayes* berbasis pembobotan *Term Frequency-Inverse Document Frequency* (TF-IDF) karena bekerja berdasarkan frekuensi kemunculan kata, sehingga lebih akurat dalam menangkap esensi sentimen pengguna [5].

Roadmap penelitian ini meliputi pengumpulan data ulasan COD Shopee, *text preprocessing*, pembobotan TF-IDF, pembagian data 80:20 (metode *hold-out*), klasifikasi dengan *Multinomial Naive Bayes*, serta evaluasi menggunakan *confusion matrix* (akurasi, *precision*, *recall*, dan *F1-score*). Penelitian ini bertujuan untuk menganalisis sentimen ulasan pengguna Shopee terhadap fitur COD. Hasilnya diharapkan menjadi bahan evaluasi peningkatan layanan bagi Shopee serta referensi akademis bagi studi analisis sentimen ke depan [6].

METODE PENELITIAN

Tahapan penelitian yang dilakukan pada penelitian ini menggunakan algoritma Naive Bayes sebagai model dalam klasifikasi yang kemudian dilakukan proses evaluasi [7]. Penjelasan secara detail mengenai tahapan metode penelitian disajikan pada Gambar 1



Gambar 1. Langkah-langkah Penelitian

Pengumpulan Data

Pengumpulan data merupakan salah satu cara untuk mencari berbagai data dalam penelitian ini [8]. Penelitian ini menggunakan data sekunder yang diperoleh dari platform Kaggle. Dataset yang digunakan berisi ulasan spesifik pengguna aplikasi Shopee mengenai fitur pembayaran COD (Cash on

Delivery) dalam bentuk teks yang diunduh dalam format CSV. Sebanyak 1.000 data ulasan pengguna digunakan sebagai objek penelitian. Spesifikasi perangkat lunak yang digunakan untuk mengolah dataset ini adalah Google Colaboratory dengan bahasa pemrograman Python, dibantu pustaka Pandas untuk manipulasi data teks dan Scikit-Learn untuk pemodelan machine learning [9]. Karakteristik dan distribusi pelabelan dari data ulasan yang digunakan dalam penelitian ini akan dijabarkan pada Tabel 1.

Tabel 1. Karakteristik Dataset Ulasan COD Shopee

Karakteristik Data	Keterangan/Jumlah
Sumber Data	Platform Kaggle (Sekunder)
Format File	Comma-Separated Values (.csv)
Total Sampel Ulasan	1.000 Data
Jumlah Sentimen Negatif	546 Data
Jumlah Sentimen Positif	454 Data
Bahasa	Bahasa Indonesia (Formal & <i>Slang</i>)

Setelah data dikumpulkan dan dipetakan karakteristiknya, data tersebut kemudian diarahkan melalui beberapa langkah penelitian yang dituliskan secara runtut dalam bentuk narasi sebagai berikut: (1) pembersihan data melalui text preprocessing, (2) transformasi teks menjadi nilai numerik dengan pembobotan TF-IDF, (3) pembagian data latih dan data uji, (4) klasifikasi sentimen menggunakan algoritma Multinomial Naive Bayes, dan (5) evaluasi performa model menggunakan confusion matrix [10]

Text Preprocessing

Text preprocessing merupakan tahapan penting untuk mengubah data ulasan mentah menjadi data yang bersih dan terstruktur sebelum dilakukan proses klasifikasi [11]. Tahapan preprocessing yang diimplementasikan pada penelitian ini meliputi: (1) case folding, yaitu proses mengubah seluruh huruf dalam teks ulasan menjadi huruf kecil (lowercase) secara seragam serta menghapus karakter angka atau simbol yang tidak diperlukan; (2) tokenization, yaitu memotong rangkaian kalimat ulasan menjadi satuan kata tunggal (token) berdasarkan spasi sekaligus memisahkan tanda baca; (3) filtering (stopword removal), yaitu menyaring kata-kata yang telah ditokenisasi untuk menghapus kata hubung atau kata umum yang tidak memiliki makna penting seperti "yang", "dan", "di", dan "ke"; (4) stemming, yaitu mengubah kata-kata berimbuhan menjadi kata dasarnya menggunakan bantuan library Sastrawi khusus bahasa Indonesia guna mengurangi variasi bentuk kata tunggal dan menyeragamkan data latih [12].

Transformasi Fitur Dan Pembagian Data

Untuk mengevaluasi performa algoritma, dataset kemudian dibagi menjadi data latih (*training data*) dan data uji (*testing data*) menggunakan metode *Hold-Out*. Pada metode ini, dataset dipisahkan dengan rasio proporsi 80:20 [13]. Berdasarkan total data yang dikunci, sebanyak 80% data (800 ulasan) digunakan sebagai data latih untuk membangun model Naive Bayes, sedangkan 20% sisanya (200 ulasan) dialokasikan sebagai data uji untuk mengevaluasi kinerja model. Pembagian dengan rasio 80:20 ini mengadopsi prinsip Pareto yang umum digunakan dalam *data science* demi menghasilkan model yang stabil dan menghindari risiko *overfitting* [14].

Klasifikasi Multinomial Naive Bayes

Multinomial Naive Bayes merupakan varian dari algoritma Naive Bayes yang dikembangkan khusus untuk menangani klasifikasi pada data tekstual, di mana fitur yang dihitung adalah frekuensi kemunculan kata dalam dokumen [15]. Algoritma ini bekerja dengan asumsi dasar bahwa setiap fitur kata bersifat independen atau tidak saling memengaruhi satu sama lain di dalam kelas yang sama. Dalam analisis sentimen ulasan Shopee, algoritma ini menghitung probabilitas posterior dari suatu dokumen ulasan untuk menentukan apakah dokumen tersebut masuk ke dalam kelas sentimen positif atau negatif. Persamaan matematis standar Multinomial Naive Bayes didefinisikan sebagai berikut:

$$P(c|d) \propto P(c) \prod_{i=1}^n P(w_i|c) \quad (1)$$

Sedangkan untuk menghitung probabilitas setiap kata (w_i) pada kelas c dengan menggunakan perataan Laplace (*Laplace Smoothing*) demi menghindari nilai probabilitas nol adalah sebagai berikut:

$$P(w_i|c) = \frac{T_{ct} + 1}{\sum_{t' \in V} T_{ct'} + |V|} \quad (2)$$

Keterangan: $P(c|d)$ adalah probabilitas posterior dari kelas c pada dokumen d ; $P(c)$ adalah probabilitas prior dari kelas c ; $P(w_i|c)$ adalah probabilitas kemunculan kata w_i pada kelas c ; T_{ct} adalah frekuensi kemunculan kata t pada kelas c ; dan $|V|$ adalah total kata unik dalam seluruh kosakata (vocabulary).

HASIL DAN PEMBAHASAN

Berdasarkan pengujian pada data ulasan pengguna Shopee terkait fitur *Cash on Delivery* (COD), diperoleh total data bersih sebanyak 1.000 ulasan. Distribusi label setelah melewati proses penyaringan otomatis menunjukkan kelas sentimen negatif sedikit lebih dominan dibandingkan sentimen positif. Detail sebaran data untuk masing-masing kelas sentimen disajikan pada Tabel 2.

Tabel 2. Distribusi Kelas Sentimen Dataset

No.	Kelas Sentimen	Jumlah Ulasan	Presentase
1	Negatif	546	54,60%
2	Positif	454	45,40%
3	Total	1.000	100,00%

Data pada Tabel 2 kemudian dibagi menggunakan teknik *Hold-Out* dengan proporsi rasio 80:20. Proses ini menghasilkan 800 data latih (*training data*) untuk pembentukan model probabilitas kata, serta 200 data uji (*testing data*) untuk mengukur performa akhir pengujian model.

Evaluasi Performa Model Klasifikasi

Pengujian performa model *Multinomial Naive Bayes* yang dikombinasikan dengan pembobotan TF-IDF menghasilkan tingkat akurasi global sebesar 78,00%. Persamaan matematis untuk menghitung nilai akurasi global pada pengujian ini disajikan sebagai berikut:

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

Berdasarkan rumus (3), hasil kalkulasi performa model dievaluasi terlebih dahulu menggunakan *Confusion Matrix* untuk melihat ketepatan prediksi pada 200 data uji. Hasil visualisasi *Confusion Matrix* dari luaran coding Python disajikan pada Gambar 2

Confusion Matrix:
 $\begin{bmatrix} 103 & 12 \\ 32 & 53 \end{bmatrix}$

Gambar 2. Hasil confusion matrix pengujian model ulasan COD

Berdasarkan acuan pada Gambar 2, diketahui bahwa model berhasil memprediksi secara tepat sebanyak 103 data sentimen negatif (*True Negative*) dan 53 data sentimen positif (*True Positive*). Sebaliknya, model mengalami misklasifikasi pada 12 data aktual negatif (*False Positive*) dan 32 data aktual positif (*False Negative*). Untuk melihat kualitas klasifikasi secara lebih mendalam, kalkulasi metrik *Precision*, *Recall*, dan *F1-Score* untuk tiap kelas dijabarkan pada Tabel 3.

Tabel 3. Distribusi Kelas Sentimen Dataset

No.	Kelas sentimen	Precision	Recall	F1 -score	Support
1	Negatif	0,76	0,90	0,82	115
2	positif	0,82	0,62	0,71	85
3	Akurasi global			0,78	200

Pembahasan Temuan Dan Analisis Fenomena

Hasil pengujian pada Tabel 3 menunjukkan adanya fenomena performa yang unik pada model Multinomial Naive Bayes. Nilai Recall untuk sentimen negatif mencapai angka tertinggi sebesar 90% (0,90). Hal ini mengindikasikan bahwa algoritma memiliki sensitivitas yang sangat tinggi dan jauh lebih andal dalam mengenali ulasan keluhan pengguna terkait fitur COD dibandingkan ulasan sentimen positif yang nilai recall-nya hanya 62%.

Tingginya kemampuan model dalam mengidentifikasi teks negatif ini didorong oleh karakteristik empiris data ulasan COD di Indonesia yang cenderung ekspresif dan repetitif saat mengajukan komplain. Isu-isu spesifik yang terjadi saat ini seperti kurir yang tidak ramah, barang rusak sebelum dibuka, salah paham sistem pembayaran, hingga kasus pembeli menolak paket, diekspresikan dengan kosakata kemunculan kata yang tegas dan berulang. Pembobotan TF-IDF berhasil memberikan nilai signifikansi yang tinggi pada pola kosakata keluhan tersebut, sehingga varian Multinomial Naive Bayes dapat menangkap frekuensi fiturnya secara optimal.

Di sisi lain, nilai precision sentimen positif yang menyentuh 82% membuktikan bahwa jika model memprediksi suatu ulasan bernilai positif, hasil prediksi tersebut memiliki tingkat kebenaran yang tinggi. Secara keseluruhan, performa akurasi 78,00% membuktikan kombinasi TF-IDF dan Multinomial Naive Bayes sangat layak diandalkan oleh manajemen e-commerce dalam memetakan dinamika sentimen fitur COD secara otomatis demi peningkatan kualitas layanan di masa mendatang.

SIMPULAN

Berdasarkan hasil analisis dan pembahasan yang telah dilakukan, dapat disimpulkan bahwa penerapan algoritma *Multinomial Naive Bayes* yang dikombinasikan dengan pembobotan *Term Frequency-Inverse Document Frequency* (TF-IDF) terbukti efektif dalam mengklasifikasikan sentimen ulasan pengguna Shopee terkait fitur *Cash on Delivery* (COD). Model klasifikasi yang dibangun menggunakan 1.000 data sampel dengan proporsi *hold-out* 80:20 ini mampu menghasilkan tingkat akurasi global yang baik, yaitu sebesar 78,00%. Karakteristik empiris data ulasan COD di Indonesia yang sarat akan repetisi kosakata emosional berhasil ditangkap secara optimal oleh model, di mana algoritma menunjukkan sensitivitas yang sangat tinggi dalam mengenali keluhan atau sentimen negatif pengguna dengan capaian nilai *recall* mencapai 90%. Evaluasi ini secara jelas menjawab bahwa dinamika sistem pembayaran COD di lapangan cenderung didominasi oleh polarisasi opini negatif akibat kendala teknis dan salah paham regulasi, namun sistem komputer mampu memetakan kecenderungan tersebut secara otomatis dengan tingkat keandalan yang tinggi.

Saran praktis yang dapat diajukan berdasarkan temuan penelitian ini ditujukan kepada dua pihak utama, yaitu pihak manajemen Shopee dan peneliti selanjutnya. Kepada manajemen Shopee, disarankan untuk memanfaatkan sistem klasifikasi otomatis ini sebagai alat pemantauan kualitas layanan (*monitoring tool*) secara *real-time*, serta segera melakukan edukasi visual yang lebih masif mengenai regulasi COD di dalam aplikasi guna menekan angka sentimen negatif yang disebabkan oleh konflik antara kurir dan pembeli di lapangan. Sementara itu, bagi peneliti selanjutnya yang ingin mendalami bidang analisis sentimen *e-commerce*, disarankan untuk memperluas cakupan penelitian dengan menambahkan tahap normalisasi kata *slang* atau bahasa gaul yang lebih komprehensif, serta mengeksplorasi penggunaan metode penyeimbangan data (*smote*) atau algoritma berbasis *deep learning* demi meningkatkan nilai *recall* pada faset sentimen positif yang masih relatif rendah dalam penelitian ini.

DAFTAR PUSTAKA

- [1] Inovasi, S. N., & Sinergi, T. 2026. "Analisis Sentimen Pengguna Shopee Terhadap Fitur COD Menggunakan Metode Naive Bayes dan Visualisasi Data. 2. <https://doi.org/10.21009/sinergi.v2i1.67688>
- [2] Filieri, Raffaele; Mcleay, Fraser; Tsui, Bruce; Lin, Zhibin *Information & Management*. 2018.
- [3] Alvandy, R. 2025. "Analisis Sentimen Ulasan E-Commerce Menggunakan Metode SVM." 18(2), 164–175.
- [4] Puspa, T., Sanjaya, R., Fauzi, A., Fitri, A., & Masruriyah, N. 2023. "Analisis sentimen ulasan pada e-commerce shopee menggunakan algoritma naive bayes dan support vector machine Analysis of review sentiment on shopee e-commerce using the naive bayes algorithm and support vector machine." 4, 16–26. <https://doi.org/10.37373/infotech.v4i1.422>
- [5] Liu, B. 2018. "Sentiment Analysis Essentials."
- [6] Puspa, T., Sanjaya, R., Fauzi, A., Fitri, A., & Masruriyah, N. 2023. "Analisis sentimen ulasan pada e-commerce shopee menggunakan algoritma naive bayes dan support vector machine Analysis of review sentiment on shopee e-commerce using the naive bayes algorithm and support vector machine." 4, 16–26. <https://doi.org/10.37373/infotech.v4i1.422>
- [7] Tan, K. L., & Lee, C. P. 2023. "applied sciences A Survey of Sentiment Analysis: Approaches, Datasets, and Future Research."
- [8] Mäntylä, M. V, Graziotin, D., & Kuuttila, M. 2018. "The Evolution of Sentiment Analysis - A Review of Research Topics, Venues, and Top Cited Papers." 27(February), 16–32.
- [9] Carneiro, T. 2018. Deep Learning for Sentiment Analysis : A Survey Article Type : Overview OPINION ADVANCED REVIEW PRIMER FOCUS ARTICLE OVERVIEW Authors : <https://doi.org/10.1002/widm.1253>

- [10] Indriyani, F. A., Fauzi, A., & Faisal, S. 2023. "Analisis sentimen aplikasi tiktok menggunakan algoritma naïve bayes dan support vector machine Tiktok application sentiment analysis using naïve bayes algorithm and support vector machine." 10, 176–184. <https://doi.org/10.37373/tekno.v10i2.419>
- [11] Nandika, M. 2025. "Implementasi Khiyar Pada Jual Beli Cash On Delivery (COD) Di Tiktok Shop Menurut Implementasi Khiyar."
- [12] Carneiro, T., Victor, R., Da, M., Nepomuceno, T., Bian, G., & Albuquerque, V. H. C. D. E. 2018. Performance Analysis of Google Colaboratory as a Tool for Accelerating Deep Learning Applications. 61677–61685. <https://doi.org/10.1109/ACCESS.2018.2874767>
- [13] Kumar, S., Pratim, P., & Nov, A. I. 2023. "A Comprehensive Review on Sentiment Analysis: Tasks, Approaches and Applications." 1–29.
- [14] Bahtiar, K. F., & Firmansyah, F. 2024. "The Influence of Online Customer Reviews and Ratings on Purchasing Decisions Through Consumer Trust as A Mediation Variable." Study of Camilan Khanza Consumers in Malang. 18(October), 695–712.
- [15] Putra, S. M., & Sundari, D. (2025). Klarifikasi Sentimen Ulasan Penggunaan Aplikasi Shopee dengan Algoritma Navis Bayes dan ID3." 9(6), 9638–9643.