



ANALISIS SENTIMEN ULASAN PRODUK SPEAKER BLUETOOTH PORTABLE DI SHOPEE MENGGUNAKAN NAIVE BAYES

Nelgi Pinando¹, Indra Lesmana², Indra Susena³, Wardianto⁴

^{1,2,3,4} Program Studi Informatika, Fakultas Sains dan Teknologi, Universitas Nurul Huda, Indonesia

Article Info

Article History:

Received: 02-January-2026

Revised: 02-March-2026

Published: 02-July-2026

Keywords:

Sentiment Analysis,

Multinomial Naive Bayes,

Bluetooth Speaker.

ABSTRACT

The e-commerce sector in Indonesia, particularly Shopee, contains vast consumer reviews for electronic products like portable Bluetooth speakers. These reviews are valuable but too numerous for manual analysis. This research applies the Multinomial Naive Bayes method to classify sentiment in these reviews. Using 3,021 data points from Kaggle, the research involved text preprocessing, rating-based labeling, TF-IDF feature extraction, and an 80:20 data split. Results show a global accuracy of 88.31%. However, the model exhibits bias toward the positive majority class, with low recall for negative and neutral classes due to data imbalance. Findings confirm text mining is effective for large-scale opinion mapping, but future research should utilize data balancing techniques or alternative algorithms to improve precision for minority classes.

Kata Kunci:

Analisis Sentimen,

Multinomial Naive Bayes,

Speaker Bluetooth.

ABSTRAK

Sektor *e-commerce* di Indonesia, khususnya Shopee, menyimpan banyak ulasan konsumen terhadap produk elektronik seperti *speaker bluetooth portable*. Ulasan tersebut bernilai tinggi namun terlalu banyak untuk dianalisis manual. Penelitian ini menerapkan metode *Multinomial Naive Bayes* untuk klasifikasi sentimen ulasan tersebut. Dengan 3.021 data dari Kaggle, tahapan penelitian meliputi praproses teks, pelabelan berbasis *rating*, ekstraksi fitur TF-IDF, dan pembagian data 80:20. Hasil pengujian menunjukkan akurasi 88,31%. Namun, model menunjukkan bias kuat terhadap kelas mayoritas positif, dengan deteksi rendah pada kelas negatif dan netral akibat ketidakseimbangan data. Temuan ini menegaskan efektivitas *text mining* untuk pemetaan opini publik, namun diperlukan teknik penyeimbangan data atau algoritma alternatif di masa depan untuk meningkatkan presisi klasifikasi kelas minoritas.

Corresponding Author

Nelgi Pinando,

Program Studi Informatika,

Universitas Nurul Huda,

nelgipinando@gmail.com

This is an open access article under the [CC BY-SA](#) license



1. INTRODUCTION

Sektor *e-commerce* di Indonesia mengalami pertumbuhan yang sangat pesat dalam beberapa tahun terakhir, dengan Shopee sebagai salah satu platform yang paling dominan digunakan oleh masyarakat. Kehadiran fitur ulasan (*review*) di platform ini menjadi sarana utama bagi konsumen untuk membagikan pengalaman nyata mereka setelah membeli sebuah produk, di mana setiap teks komentar tersebut menyimpan informasi mengenai pengalaman pengguna terhadap kualitas produk [1]. Secara spesifik, pada kategori produk elektronik seperti *speaker bluetooth portable*, konsumen memiliki ekspektasi yang sangat tinggi terhadap berbagai aspek fungsionalitas, mulai dari kejernihan kualitas audio, kestabilan konektivitas nirkabel, hingga daya tahan baterai [2]. Mengingat aspek-aspek tersebut sering kali dinilai secara subjektif, ulasan-ulasan ini menjadi bahan pertimbangan yang sangat penting bagi calon pembeli dalam pengambilan keputusan sebelum bertransaksi [3]. Penumpukan data ulasan yang begitu cepat sebenarnya menyimpan informasi berharga mengenai kepuasan konsumen, yang jika diolah dengan baik, dapat memberikan gambaran objektif mengenai reputasi serta kualitas produk tersebut di mata publik. Penggunaan analisis sentimen dalam ranah *e-commerce* telah didukung oleh

berbagai temuan mutakhir yang menyoroti efektivitas algoritma pembelajaran mesin dalam memproses data ulasan konsumen berskala besar.

Namun, volume ulasan yang bisa mencapai ratusan hingga ribuan teks pada satu halaman produk membuat proses evaluasi dan pembacaan secara manual menjadi sangat tidak praktis serta memakan banyak waktu [4]. Di sinilah pendekatan *text mining* melalui analisis sentimen hadir sebagai solusi untuk mengekstrak dan mengelompokkan data teks acak tersebut menjadi informasi yang lebih terstruktur [5]. Lewat sistem ini, ulasan konsumen dapat diklasifikasikan secara otomatis ke dalam kategori opini positif maupun negatif. Proses klasifikasi otomatis ini tidak hanya membantu calon pembeli dalam memilih informasi secara instan, tetapi juga menjadi alat evaluasi yang krusial bagi para penjual atau produsen untuk memahami keluhan utama pelanggan dengan cepat demi perbaikan kualitas produk ke depan [6].

Meskipun penelitian analisis sentimen menggunakan *Naive Bayes* pada data Shopee sudah sering dilakukan, eksplorasi mendalam yang berfokus pada perangkat keras elektronik atau gawai masih terus memerlukan pembaruan. Ulasan untuk produk perangkat audio seperti *speaker* memiliki keunikan tersendiri karena konsumen sangat kritis terhadap detail performa fisik produk yang sering ditulis menggunakan perpaduan bahasa formal dan istilah kasual masyarakat, seperti kata "nge-bass", "cempreng", "semer", atau kendala sinkronisasi *Bluetooth*[7]. Dalam konteks pengolahan data tersebut, tantangan utama yang dihadapi adalah bagaimana sistem mampu memahami ambiguitas makna dalam bahasa sehari-hari yang sering kali tidak tercatat dalam kamus standar. Oleh karena itu, pemilihan metode yang tepat menjadi kunci dalam memastikan bahwa setiap ulasan dapat diterjemahkan menjadi parameter kualitas yang akurat dan dapat diandalkan oleh para pemangku kepentingan. Keandalan algoritma dalam menangani kompleksitas semantik ini sangat krusial untuk memastikan bahwa hasil analisis tidak hanya sekedar angka, tetapi benar-benar mencerminkan pengalaman nyata pengguna di lapangan. Berdasarkan kondisi tersebut, penelitian ini bertujuan untuk menerapkan metode *Naive Bayes* dalam mengklasifikasikan sentimen ulasan khusus produk *speaker bluetooth portable* di Shopee [8]. Melalui pendekatan ini, diharapkan hasil analisis dapat memberikan *insight* yang lebih relevan dan presisi, baik sebagai panduan belanja bagi calon konsumen maupun bahan evaluasi digital bagi produsen audio [9].

Pendekatan yang diterapkan dalam riset ini mengacu pada prinsip pengembangan sistem berbasis data yang berfokus pada kualitas representasi fitur linguistik dalam teks. Pemilihan metode Multinomial Naive Bayes didasarkan pada karakteristiknya yang efisien dalam memproses dataset teks berdimensi tinggi, di mana setiap fitur kata diasumsikan independen satu sama lain dalam memberikan kontribusi terhadap probabilitas kelas sentimen [10][11]. Selain itu, integrasi teknik *text mining* yang komprehensif bertujuan untuk menanggulangi variasi bahasa yang tidak baku pada ulasan e-commerce, sehingga model mampu melakukan generalisasi dengan lebih akurat [12][13]. Untuk mengantisipasi tantangan ketidakseimbangan kelas (data imbalance) yang umum ditemukan pada ulasan produk, penelitian ini juga mengacu pada standar evaluasi berbasis metrik yang ketat guna memastikan bahwa hasil klasifikasi tidak hanya mengandalkan akurasi global, namun juga memberikan perhatian pada performa model dalam mengenali kelas minoritas [14][15].

2. RESEARCH METHOD

Penelitian ini menggunakan pendekatan kuantitatif dengan menerapkan metode Analisis Sentimen berbasis teknik *text mining* untuk melakukan klasifikasi otomatis terhadap dokumen ulasan pengguna. Kerangka kerja penelitian ini dirancang secara sistematis agar data ulasan yang bersifat tidak terstruktur (acak) dapat ditransformasikan ke dalam matriks numerik yang valid untuk diolah menggunakan perhitungan probabilitas algoritma *Multinomial Naive Bayes*. Seluruh tahapan riset, mulai dari perancangan konsep, akuisisi data, pengolahan, hingga evaluasi performa, dilakukan dengan merujuk pada standar prosedur pemrosesan bahasa alami yang dapat dipertanggungjawabkan secara ilmiah. Melalui tahapan yang terukur, penelitian ini bertujuan menghasilkan model klasifikasi yang mampu memetakan polaritas sentimen dengan tingkat akurasi yang objektif, sehingga memberikan wawasan strategis yang dapat diaplikasikan dalam pengambilan keputusan berbasis data.

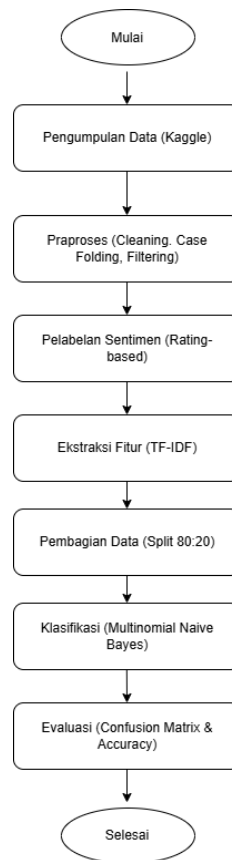
Penggunaan teknik analisis sentimen berbasis teks ini didukung oleh berbagai temuan mutakhir yang menunjukkan bahwa algoritma pembelajaran mesin sangat efektif dalam memproses data ulasan konsumen berskala besar. Penelitian-penelitian tersebut secara konsisten membuktikan bahwa penerapan metode klasifikasi teks mampu mengidentifikasi polaritas opini pengguna dengan tingkat akurasi yang tinggi, sehingga menjadikannya instrumen krusial dalam analisis data e-commerce [16][17][18].

Berikut adalah rincian tahapan sistematis dalam penelitian ini:

1. Alur Penelitian

Alur penelitian ini dilakukan secara bertahap untuk mengolah data teks ulasan menjadi prediksi sentimen yang akurat. Proses dimulai dari Pengumpulan Data mentah, dilanjutkan dengan Praproses Teks (membersihkan simbol, menyeragamkan huruf kecil, dan merapikan spasi), serta Pelabelan Sentimen otomatis berdasarkan nilai *rating* bintang.

Setelah data bersih dan berlabel, dilakukan Ekstraksi Fitur (TF-IDF) untuk mengubah teks menjadi angka, diikuti Pembagian Data (Split 80:20) menjadi data latih dan data uji. Terakhir, data dilatih menggunakan algoritma Klasifikasi Multinomial Naive Bayes dan diuji pada tahap Evaluasi menggunakan *Confusion Matrix* serta tingkat *Accuracy* untuk mengukur performa akhir model.



Gambar 1 Diagram Alur Penelitian

2. Pengumpulan Data (*Data Acquisition*)

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari repositori data publik Kaggle. Pemilihan dataset dari platform ini didasarkan pada ketersediaan variabel yang lengkap, seperti teks ulasan (*review*), rating produk, dan atribut metadata lainnya yang diperlukan untuk analisis sentimen. Penggunaan dataset dari Kaggle memberikan keuntungan berupa efisiensi waktu dalam fase akuisisi data dan memastikan bahwa dataset yang digunakan memiliki volume yang cukup besar untuk melatih model klasifikasi secara optimal. Setelah diunduh, data tersebut diproses untuk memastikan format file sesuai dengan kebutuhan sistem, yaitu dalam format .xlsx atau .csv, guna meminimalisir kendala kompatibilitas selama tahap pemrosesan awal[19].

3. Pembersihan Dan Praproses Data (*Data Cleaning & Preprocessing*)

Tahap ini mengintegrasikan pembersihan data dan pemrosesan teks ke dalam satu alur kerja komputasi yang terpadu. Pembersihan data dimulai dengan menghapus baris kosong (*drop null values*) serta menghilangkan data duplikat (*drop duplicates*) untuk memastikan setiap ulasan adalah entri yang unik dan objektif. Selanjutnya, dilakukan proses praproses yang meliputi:

- (a) *Case folding*
Untuk mengubah seluruh karakter menjadi huruf kecil
- (b) *Noise filtering*
Menghapus semua karakter non-alfanumerik (tanda baca, simbol, angka) guna membuang elemen yang tidak berkontribusi pada makna sentimen
- (c) *Stripping*
Menghapus spasi berlebih antar kata agar struktur teks menjadi rapi dan konsisten. Integrasi langkah ini sangat krusial untuk menyeragamkan format teks serta meminimalkan ambiguitas kata sebelum masuk ke tahap ekstraksi fitur.

4. Pelabelan Sentimen (*Sentiment Labeling*)

Data hasil pembersihan dan praproses dikategorikan berdasarkan rating bintang pembeli (Positif, Netral, Negatif). Label ini kemudian ditransformasikan ke dalam representasi numerik (indeks angka) agar dapat diproses oleh fungsi matematis algoritma.

5. Ekstraksi Fitur (*Feature Extraction – TF-IDF*)

Data teks yang sudah bersih diubah menjadi vektor numerik melalui metode *Term Frequency-Inverse Document Frequency* (TF-IDF). Metode ini memberikan bobot pada setiap kata berdasarkan frekuensi kemunculannya dalam sebuah ulasan dibandingkan dengan tingkat kelangkaannya dalam seluruh dataset [6][7].

6. Pembagian Data (*Data Splitting*)

Dataset dibagi menjadi dua bagian dengan proporsi 80% untuk *training set* dan 20% untuk *testing set*. Pembagian acak dengan parameter *stratify* diterapkan untuk memastikan distribusi kelas sentimen yang seimbang, sehingga model memiliki basis pembelajaran yang representatif [2][5].

7. Pemodelan Klasifikasi (*Multinomial Naive Bayes*)

Model dibangun menggunakan algoritma *Multinomial Naive Bayes* (MNB). Algoritma ini bekerja dengan menghitung peluang probabilitas bersyarat dari fitur kata hasil TF-IDF untuk menentukan prediksi kelas sentimen dengan nilai probabilitas tertinggi [4].

8. Penerapan Metode Multinomial Naïve Bayes

Untuk memberikan gambaran prosedural mengenai tahapan komputasi yang dilakukan secara teknis, berikut adalah algoritma dalam bentuk *pseudocode* yang digunakan dalam membangun model klasifikasi:

Algoritma Klasifikasi Sentimen Ulasan:

1. INPUT: Dataset_Raw (.xlsx)
2. PROSES Data_Cleaning: (Hapus null/duplikat, Case folding, Noise filtering, Stripping)
3. PROSES Pelabelan: (Konversi rating ke label Positif/Netral/Negatif & Encoding 0-2)
4. PROSES Ekstraksi_Fitur: (Transformasi teks ke matriks numerik menggunakan TF-IDF)
5. PROSES Pembagian_Data: (Split data 80% Training dan 20% Testing)
6. PROSES Pelatihan_Model: (Inisialisasi & Fit model Multinomial Naive Bayes)
7. PROSES Evaluasi: (Prediksi testing & Perhitungan Accuracy, Precision, Recall, F1)
8. OUTPUT: Laporan performa model dan hasil prediksi

3. RESULTS AND ANALYSIS

Bagian ini menjelaskan hasil penelitian serta memberikan pembahasan yang komprehensif terkait efektivitas algoritma Naïve Bayes dalam melakukan analisis sentimen terhadap ulasan produk speaker bluetooth pada platform Shopee. Hasil penelitian disajikan dalam bentuk tabel dan metrik evaluasi untuk memudahkan pembaca dalam memahami temuan utama [9].

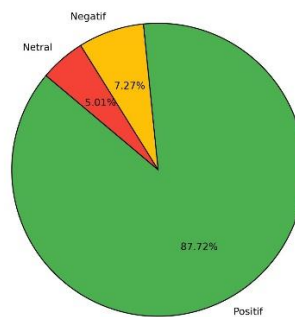
3.1. Dataset dan Preprocessing (*Quattrocento Sans, BOLD, Capitalize Each Word, 10 PT*)

Penelitian ini menggunakan dataset ulasan produk speaker bluetooth yang bersumber dari kaggle. Tahap awal penelitian dimulai dengan pengumpulan 3.021 data ulasan mentah. Guna memastikan kualitas data yang akan diklasifikasikan, dilakukan tahapan *preprocessing*. Proses ini krusial untuk menghilangkan teks yang tidak memiliki nilai informasi semantik, sehingga model dapat fokus pada fitur yang relevan [10]. Sebagai gambaran mengenai perubahan teks selama proses *preprocessing*, Tabel 2 menyajikan contoh transformasi data untuk setiap kategori sentimen.

Table 1 Transformasi Data dan Pelabelan Sentimen

Kategori	Label	Ulasan Mentah	Hasil Cleaning
Negatif	0	Fitur Terbaik:salon kualitas buruk banget rusak Sepadan dengan Harga:harganya agak mahal Salon rusak jangan di jual woi kasihan yg beli kalau gini gmn ini kualitas ajor kecewa aku	fitur terbaik salon kualitas buruk banget rusak sepadan dengan harga harganya agak mahal salon rusak jangan di jual woi kasihan yg beli kalau gini gmn ini kualitas ajor kecewa aku kalau bisa jangan beli di toko ini lagi ya gak amanah sm sekali

		.kalau bisa jangan beli di toko ini lagi ya gak amanah sm sekali..	
Netral	1	Maaf sy kasih bintang 3 „speaker blutut nya cuma bisa dipake micro SD sm mic nya doang,,perangkat blutut nya ga konek,,sedikit kecewa,,ya maklum karena harga nya juga mungkin jadinya begini.TKS	maaf sy kasih bintang 3 speaker blutut nya cuma bisa dipake micro sd sm mic nya doang perangkat blutut nya ga konek sedikit kecewa ya maklum karena harga nya juga mungkin jadinya begini tks
Positif	2	Sepadan dengan Harga:worth it! Fitur Terbaik:Lucu bgt mic nya kecil, ringan Gampang banget connect ke bluetooth, nyaring suaranya. Hahaha suka sih jadi hemat bisa karaokean dimanapun, kapapun. Ringan bgt, bisa ditenteng dibawa kemanapun. Terimakasih 🌟👍	sepadan dengan harga worth it fitur terbaik lucu bgt mic nya kecil ringan gampang banget connect ke bluetooth nyaring suaranya hahaha suka sih jadi hemat bisa karaokean dimanapun kapapun ringan bgt bisa ditenteng dibawa kemanapun terimakasih



Gambar 2 Persentase Distribusi Kelas Sentimen

Gambar 2 menunjukkan persentase distribusi kelas sentimen pada dataset. Terlihat adanya ketidakseimbangan data (*data imbalance*) yang ekstrem, di mana sentimen Positif sangat mendominasi dengan 87.69%, diikuti Negatif sebesar 7.27%, dan Netral sebesar 5.01%.

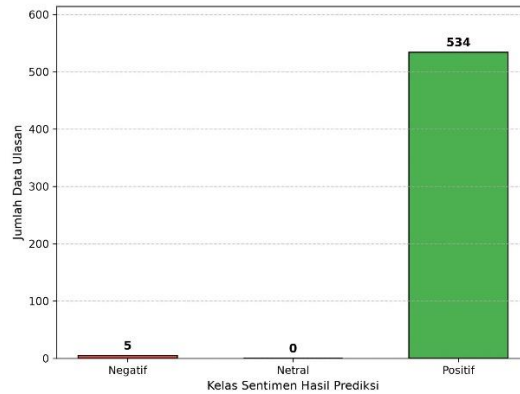
3.2. Prosedur Pembagian Data (*Quattrocento Sans, BOLD, Capitalize Each Word, 10 PT*)

Data yang telah bersih kemudian dibagi menjadi data latih (*training*) dan data uji (*testing*) dengan rasio 80:20 untuk memastikan model dievaluasi pada data yang belum pernah dipelajari sebelumnya [4]. Rincian pembagian data tersebut disajikan pada Tabel 3.

Table 2 Data Split Distribution

Description	Number Of Data	Percentage
Training Data	2.157	80 %
Testing Data	539	20 %
Total	2.696	100 %

Berdasarkan Tabel 3, Sebanyak 2.157 data digunakan untuk membangun pola model, sedangkan 539 data sisanya digunakan sebagai data uji untuk mengevaluasi kemampuan generalisasi model Naïve Bayes. Hasil klasifikasi pada 539 data uji tersebut menunjukkan tingkat akurasi sebesar 88,31%.

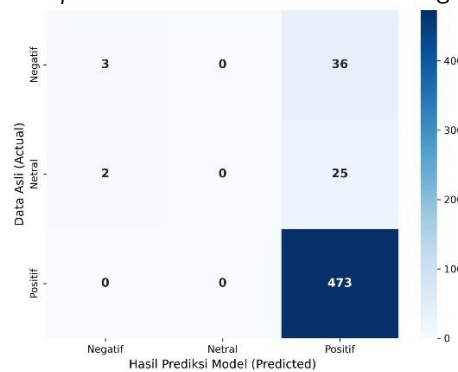


Gambar 3 Grafik Hasil Akhir Klasifikasi Prediksi Naive Bayes

Visualisasi hasil akhir prediksi model terhadap data uji (*testing*) disajikan pada **Gambar 3**. Grafik batang tersebut memperlihatkan dominasi klasifikasi yang sangat kuat pada kelas Positif dengan total mencapai 534 data ulasan. Sebaliknya, model tampak mengalami kesulitan signifikan dalam mengenali karakteristik linguistik pada kelas minoritas, yang terlihat dari sangat rendahnya jumlah prediksi untuk kelas Negatif (5 data) dan kelas Netral yang sama sekali tidak mendapatkan hasil prediksi (0 data).

3.3. Analisis Confusion Matrix

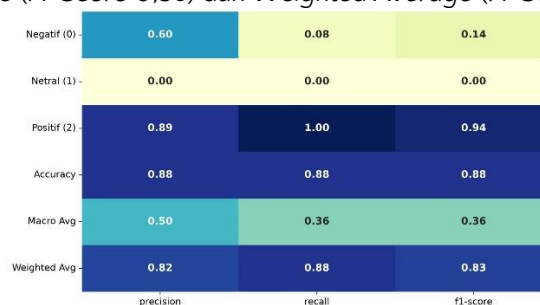
Berdasarkan hasil eksekusi program, *confusion matrix* model adalah sebagai berikut:



Gambar 4 Confusion Matrix

3.4. Pembahasan dan Analisis Mendalam

Analisis terhadap performa model menunjukkan adanya bias yang kuat terhadap kelas mayoritas, yakni Kelas 2, yang terlihat dari capaian *recall* sebesar 1,00. Fenomena ini sejalan dengan penelitian Mubarak et al. yang menyatakan bahwa algoritma Naïve Bayes cenderung memprediksi kelas mayoritas secara dominan ketika distribusi data antar kelas sangat timpang, sehingga memicu terjadinya *over-prediction* di mana model cenderung mengelompokkan data Kelas 0 dan Kelas 1 ke dalam Kelas 2. Kondisi ini dipertegas dengan adanya perbedaan signifikan antara *Macro Average (F1-Score 0,36)* dan *Weighted Average (F1-Score 0,83)*



Gambar 5 Classification Report Naive Bayes

Pada Gambar 5 di mana nilai *weighted average* yang tinggi mencerminkan dominasi akurasi pada kelas mayoritas, sementara *macro average* yang rendah secara jujur menunjukkan kegagalan model dalam mengenali kelas minoritas (Kelas 0 dan 1) secara adil.

Kegagalan pada Kelas 1 dengan *F1-Score* 0,00 mengindikasikan bahwa fitur yang diekstraksi dari data latih belum mampu membedakan karakteristik linguistik Kelas 1 dibandingkan dengan Kelas 2, yang merupakan tantangan umum dalam klasifikasi teks dengan ketidakseimbangan data (*imbalanced data*) di mana model cenderung mengabaikan kelas minoritas demi memaksimalkan tingkat akurasi global. Kendati demikian, pencapaian akurasi global sebesar 88,31% membuktikan efisiensi komputasi Naïve Bayes yang tetap kompetitif untuk klasifikasi teks berskala besar di platform e-commerce dibandingkan dengan model *deep learning* yang memerlukan sumber daya komputasi tinggi.

4. CONCLUSION

Penelitian ini berhasil menerapkan metode Multinomial Naive Bayes sebagai kerangka kerja utama untuk melakukan analisis sentimen terhadap ulasan produk speaker *bluetooth portable* pada platform Shopee. Proses penelitian ini dilakukan secara sistematis, dimulai dari akuisisi data ulasan konsumen, dilanjutkan dengan serangkaian tahap praproses teks yang ketat, meliputi pembersihan data (*data cleaning*), *case folding*, serta normalisasi teks untuk meminimalkan *noise* linguistik. Data yang telah bersih kemudian diekstraksi menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF) untuk mentransformasikan data tekstual yang bersifat tidak terstruktur menjadi representasi numerik yang valid dan informatif, sehingga dapat diproses secara optimal oleh algoritma klasifikasi.

Berdasarkan hasil pengujian eksperimental, algoritma Multinomial Naive Bayes terbukti mampu mengklasifikasikan ulasan ke dalam kategori sentimen positif, netral, dan negatif dengan performa yang cukup baik, serta mampu menjawab ekspektasi riset yang telah ditetapkan pada bab pendahuluan mengenai efektivitas *text mining* dalam memetakan pola opini publik. Hasil analisis mendalam menunjukkan bahwa mayoritas ulasan konsumen terhadap produk speaker *bluetooth portable* di Shopee memiliki kecenderungan sentimen positif, yang sekaligus membuktikan bahwa teknik *text mining* dan analisis sentimen merupakan instrumen krusial bagi calon konsumen dalam memahami kualitas produk berdasarkan pengalaman nyata pengguna lain, serta menjadi alat evaluasi digital yang strategis bagi penjual dalam memantau tingkat kepuasan pelanggan secara objektif.

Meskipun model telah mencapai tingkat akurasi yang kompetitif, penelitian ini juga mengidentifikasi adanya tantangan teknis terkait ketidakseimbangan kelas (*imbalanced data*) yang memengaruhi sensitivitas model pada kelas minoritas. Oleh karena itu, terdapat prospek pengembangan yang luas dalam riset ini di masa depan. Sebagai langkah optimalisasi, metode Naive Bayes dapat dibandingkan secara empiris dengan algoritma klasifikasi lain yang memiliki karakteristik berbeda, seperti *Support Vector Machine* (SVM) atau *Random Forest*, guna mengevaluasi stabilitas model. Selain itu, integrasi metode *Deep Learning* seperti *Long Short-Term Memory* (LSTM) atau penerapan teknik penyeimbangan data seperti SMOTE (*Synthetic Minority Over-sampling Technique*) sangat direkomendasikan untuk diterapkan pada studi selanjutnya, guna menangkap kompleksitas semantik yang lebih dalam serta meningkatkan presisi model dalam mengenali sentimen negatif dan netral secara lebih adil.

THANK-YOU NOTE

Penulis mengucapkan terima kasih kepada Program Studi Informatika Universitas Nurul Huda atas dukungan dan fasilitas yang diberikan selama proses penelitian berlangsung. Penulis juga menyampaikan apresiasi kepada semua pihak yang telah membantu dalam proses pengumpulan data, pengolahan data, hingga penyusunan artikel ilmiah ini.

REFERENCE (QUATTROCENTO SANS, BOLD, UPPERCASE, 10 PT)

The primary reference is an international journal article. All references must be to the most relevant and up-to-date source, with a minimum of 15 references. References must be written using Mendeley in IEEE style. Please use a consistent format for references.

For Example :

- [1] E. H. Muktafin, Kusriani, dan E. T. Luthfi, "Analisis Sentimen pada Ulasan Pembelian Produk di Marketplace Shopee Menggunakan Pendekatan Natural Language Processing," *J. Eksplora Inform.*, vol. 10, no. 1, 2020.
- [2] S. Mubarak, T. Harmini, dan F. R. Pradana, "Analisis Sentimen Ulasan Produk Elektronik pada E-Commerce Menggunakan Metode Naïve Bayes," *Pros. Semin. Nas. Sains dan Teknol.*, vol. 3, no. 1, 2026.

-
- [3] A. Nurian, M. S. Ma'arif, I. N. Amalia, dan C. Rozikin, "Analisis Sentimen Pengguna Aplikasi Shopee pada Situs Google Play Menggunakan Naive Bayes Classifier," *JITET (Jurnal Inform. dan Tek. Elektro Terapan)*, vol. 12, no. 1, 2024.
- [4] A. H. Hasugian, M. Fakhriza, dan D. Zukhoiriyah, "Analisis Sentimen Pada Review Pengguna E-Commerce Menggunakan Algoritma Naive Bayes," *J. Teknol. Sist. Inf. dan Sist. Komput. TGD*, vol. 6, no. 1, hlm. 98–107, 2023.
- [5] R. L. Simanjuntak, T. R. Siagian, V. Anggriani, dan Arnita, "Analisis Sentimen Ulasan Pada Aplikasi E-Commerce Shopee Dengan Menggunakan Algoritma Naive Bayes," *J. Tek. Mesin, Elektro dan Ilmu Komput. (TEKNIK)*, vol. 3, no. 3, hlm. 23–39, 2023.
- [6] T. P. R. Sanjaya, A. Fauzi, dan A. F. N. Masruriyah, "Analisis sentimen ulasan pada e-commerce shopee menggunakan algoritma naive bayes dan support vector machine," *INFOTECH: J. Inform. Teknol.*, vol. 4, no. 1, hlm. 16–26, 2023.
- [7] N. F. Octavia dan Berlilana, "Penerapan Algoritma K-Nearest Neighbor untuk Analisis Sentimen Ulasan Produk Elektronik pada Platform E-Commerce," *J. Algoritma*, vol. 22, no. 2, 2025.
- [8] C. H. Yutika, Adiwijaya, dan S. A. Faraby, "Analisis Sentimen Berbasis Aspek pada Review Female Daily Menggunakan TF-IDF dan Naive Bayes," *J. Media Inform. Budidarma*, vol. 5, no. 2, hlm. 422–430, 2021.
- [9] N. A. C. Agustina, D. H. Citra, W. Purnama, C. Nisa, dan A. R. Kurnia, "Implementasi Algoritma Naive Bayes untuk Analisis Sentimen Ulasan Shopee pada Google Play Store," *MALCOM: Indones. J. Mach. Learn. Comput. Sci.*, vol. 2, no. 1, hlm. 47–54, 2022.
- [10] A. Muzaki, V. Febriana, dan W. N. Cholifah, "Analisis Sentimen Pada Ulasan Produk Di E-Commerce Dengan Metode Naive Bayes," *J. Ris. dan Apl. Mhs. Inform.*, vol. 5, no. 4, 2024.
- [11] A. H. Hasugian, M. Fakhriza, dan D. Zukhoiriyah, "Analisis Sentimen Pada Review Pengguna E-Commerce Menggunakan Algoritma Naive Bayes," *J. Teknol. Sist. Inf. dan Sist. Komput. TGD*, vol. 6, no. 1, hlm. 98–107, 2023.
- [12] P. Maxmiliano, Y. F. Riti, I. Y. Nugroho, dan C. E. C. Juniarto, "Perbandingan Algoritma Naive Bayes dan Bert untuk Analisis Sentimen Ulasan Produk Shopee Berdasarkan Rating dan Atribut Produk (Warna/Kategori)," *J. Media Inform. [JUMIN]*, vol. 6, no. 5, hlm. 2552–2565, 2025.
- [13] M. S. Wulandari, R. Noveandini, dan N. D. Putra, "Analisis Sentimen Terhadap Ulasan Produk Pada Sistem Penjualan Toko Putra Elektronik," *J. Islamic Bus. Manag. Stud.*, vol. 2, no. 2, hlm. 84–98, 2021.
- [14] R. Taufikurohman dan S. N. Rohman, "Analisis Sentimen Ulasan Produk Smartphone Pada Platform Shopee Menggunakan Metode Naive Bayes," *Pros. Semin. Nas. Teknol. dan Sains*, vol. 5, 2026.
- [15] C. Cahyaningtyas, Y. Nataliani, dan I. R. Widiyari, "Analisis sentimen pada rating aplikasi Shopee menggunakan metode Decision Tree berbasis SMOTE," *AITI: J. Teknol. Inf.*, vol. 18, no. 2, hlm. 173–184, 2021.
- [16] A. Patel, P. Oza, and S. Agrawal, "Sentiment analysis of customer feedback and reviews for airline services using language representation model," *Procedia Computer Science*, vol. 218, pp. 2459–2467, 2023.
- [17] L. Ashbaugh and Y. Zhang, "A comparative study of sentiment analysis on customer reviews using machine learning and deep learning," *Computers*, vol. 13, no. 12, p. 340, Dec. 2024.
- [18] M. Haroon, Z. Alam, R. Kousar, J. Ahmad, and F. Nasim, "Sentiment analysis of customer reviews on e-commerce platforms: A machine learning approach," *Bulletin of Business and Economics*, vol. 13, no. 3, pp. 230–238, 2024.
- [19] N. Hadi dan D. Sugiarto, "Analisis Sentimen Pembangunan IKN pada Media Sosial X Menggunakan Algoritma SVM, Logistic Regression dan Naive Bayes," *Jurnal Informatika: Jurnal Pengembangan IT*, vol. 10, no. 1, Jan. 2025.
-