



Perbandingan Naïve Bayes dan Logistic Regression untuk Analisis Sentimen Ulasan Produk Skincare di Tokopedia

Indra Dikusuma ^{1*}, dan Yunus Widjaja ²

¹ Universitas Pembangunan Jaya; Tangerang Selatan, Banten; e-mail : indra.dikusuma@student.upj.ac.id

² Universitas Pembangunan Jaya, Tangerang Selatan, Banten; e-mail : yunus.widjaja@upj.ac.id

* Corresponding Author : Indra Dikusuma

Abstract: This study aims to compare the performance of Naïve Bayes (NB) and Logistic Regression (LR) in classifying the sentiment of skincare product reviews on Tokopedia. A total of 973 reviews were collected and processed through text preprocessing and feature extraction using TF-IDF. Evaluation using accuracy, precision, recall, and F1-score showed that NB achieved an accuracy of 75.9% but struggled to detect negative sentiment, while LR reached an accuracy of 92.3% with a more balanced classification. It is concluded that Logistic Regression is superior and more suitable for complex e-commerce review datasets, offering practical guidance for effective model selection

Keywords: Sentiment Analysis; E-commerce; Naïve Bayes; Logistic Regression; Tokopedia

Abstrak: Penelitian ini bertujuan membandingkan kinerja Naïve Bayes (NB) dan Logistic Regression (LR) dalam klasifikasi sentimen ulasan produk skincare di Tokopedia. Sebanyak 973 ulasan diproses melalui text preprocessing dan ekstraksi fitur menggunakan TF-IDF. Evaluasi menggunakan akurasi, precision, recall, dan F1-score menunjukkan NB memperoleh akurasi 75,9% namun kurang seimbang dalam mendekati sentimen negatif, sementara LR mencapai akurasi 92,3% dengan klasifikasi yang stabil. Disimpulkan bahwa Logistic Regression lebih unggul dan sesuai untuk dataset ulasan e-commerce yang kompleks, memberikan rekomendasi praktis untuk pemilihan model yang efektif.

Kata kunci: Analisis Sentimen; E-commerce; Naïve Bayes; Logistic Regression; Tokopedia

1. Pendahuluan

Perkembangan teknologi informasi dan komunikasi yang pesat telah mengubah lanskap bisnis dan pola konsumsi masyarakat secara signifikan [1], [2]. Salah satu manifestasi paling menonjol dari transformasi ini adalah kemunculan dan pertumbuhan electronic commerce (e-commerce), yang telah menjadi bagian tak terpisahkan dari gaya hidup modern di Indonesia [3].

Platform seperti Tokopedia, Shopee, dan Lazada telah menjadi arena utama bagi jutaan pengguna untuk melakukan transaksi jual beli beragam produk dan jasa [4]. Sejarah e-commerce di Indonesia sendiri dimulai sejak tahun 1999 dan terus berkembang hingga kini [1].

Dalam ekosistem e-commerce yang sangat kompetitif, ulasan pelanggan (customer review) memegang peranan krusial [5]. Ulasan menyediakan wawasan langsung mengenai pengalaman konsumen, yang menjadi sumber informasi utama bagi calon pembeli lainnya [6]. Di lingkungan belanja online, di mana konsumen tidak dapat berinteraksi fisik dengan produk, mereka sangat bergantung pada ulasan untuk menilai kualitas dan performa produk [7]. Ulasan positif terbukti secara signifikan meningkatkan kepercayaan konsumen (consumer trust) dan persepsi nilai (perceived value), yang pada gilirannya mendorong keputusan pembelian [8].

Mengingat volume ulasan yang terus bertambah, analisis manual untuk mengekstrak sentimen menjadi tugas yang tidak efisien [9]. Untuk mengatasi tantangan ini, analisis sentimen otomatis menjadi solusi yang sangat diperlukan.

Received: September 21, 2025

Revised: November 4, 2025

Accepted: November 26, 2025

Published: November 29, 2025

Curr. Ver.: November 29, 2025



Copyright: © 2025 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY SA) license (<https://creativecommons.org/licenses/by-sa/4.0/>).

Analisis sentimen, atau opinion mining, adalah studi yang menganalisis pendapat dan kesan orang yang diungkapkan dalam bentuk teks [10]. Dengan menerapkan teknik Natural Language Processing (NLP) dan machine learning, perusahaan dapat mengklasifikasikan sentimen dalam ulasan ke dalam kategori positif, negatif, atau netral secara efisien [11]. Analisis ini memungkinkan pelaku bisnis untuk memahami preferensi pelanggan dan meningkatkan kualitas layanan mereka [10].

Berbagai penelitian telah mengaplikasikan algoritma machine learning untuk melakukan analisis sentimen pada ulasan di platform e-commerce Indonesia. Tiga metode klasifikasi yang paling umum digunakan adalah Naive Bayes, Support Vector Machine (SVM), dan Regresi Logistik [3]. Studi-studi ini menunjukkan bahwa tidak ada satu pun algoritma yang unggul secara universal; kinerjanya sangat bergantung pada karakteristik dataset dan metode preprosesan data yang digunakan [12], [13].

Penelitian ini akan mensintesis temuan dari berbagai studi relevan untuk memberikan gambaran komprehensif mengenai penerapan analisis sentimen di e-commerce Indonesia, membandingkan efektivitas berbagai pendekatan [14], serta menegaskan peran vitalnya dalam dinamika pasar digital saat ini [15].

2. Kajian Pustaka atau Penelitian Terkait

Analisis sentimen ulasan produk di platform e-commerce sangat penting untuk memahami persepsi konsumen terhadap produk dan layanan. Akurasi dalam klasifikasi sentimen sangat berpengaruh terhadap pengambilan keputusan yang berbasis data, baik oleh pengembang aplikasi maupun pihak-pihak terkait yang terlibat dalam distribusi dan pemasaran produk.

Dalam beberapa tahun terakhir, metode klasifikasi sentimen telah berkembang pesat, mulai dari pendekatan statistik klasik seperti Naïve Bayes (NB) hingga metode machine learning yang lebih kompleks seperti Logistic Regression (LR). Bab ini mengulas penelitian terkait yang membahas penggunaan kedua metode tersebut dalam analisis sentimen, serta landasan teori yang mendasari pemilihan metode dalam studi ini.

2.1. Penelitian Terkait

Berbagai penelitian telah mengeksplorasi penggunaan algoritma Naïve Bayes (NB) dan Logistic Regression (LR) untuk analisis sentimen ulasan produk, baik di platform e-commerce maupun di sektor lainnya. Naïve Bayes, dengan keunggulannya dalam menangani dataset besar secara cepat, sering kali digunakan pada analisis teks yang membutuhkan klasifikasi biner, seperti analisis sentimen pada ulasan produk. Salah satu penelitian yang relevan adalah oleh Indra Kurniawan, yang membandingkan Naïve Bayes dengan Support Vector Machine (SVM) dalam klasifikasi sentimen pada ulasan produk di Tokopedia, Shopee, dan Lazada, menunjukkan bahwa Naïve Bayes dan SVM menghasilkan akurasi lebih dari 80%. Di sisi lain, Azhar et al. (2020) membandingkan Naïve Bayes dengan SVM dalam analisis sentimen, di mana Naïve Bayes memperoleh akurasi 69,74% yang meningkat menjadi 75,21% setelah optimasi fitur menggunakan Particle Swarm Optimization (PSO) [16]. Penelitian lain juga menunjukkan bahwa Naïve Bayes efektif dalam klasifikasi sentimen dengan dataset seimbang dari ulasan aplikasi di Shopee [16].

Pada sisi lain, Logistic Regression (LR) juga digunakan secara luas dalam analisis sentimen karena kemampuannya dalam menangani data yang lebih kompleks dan variabel yang saling berkorelasi. Logistic Regression mengubah fungsi linier menjadi probabilitas kelas, yang sering kali memberikan hasil yang lebih baik dalam situasi di mana hubungan antar fitur lebih rumit. Begitu juga dengan penelitian lainnya yang menunjukkan bahwa Logistic Regression lebih adaptif terhadap data yang lebih kompleks dibandingkan dengan Naïve Bayes [17].

2.2 Landasan Teori

Pemilihan metode dalam penelitian ini didasarkan pada kekuatan teoretis masing-masing model dalam menangani karakteristik data deret waktu.

a. *Naïve Bayes (NB)*

Naïve Bayes adalah algoritma klasifikasi probabilistik yang berbasis pada Teorema Bayes, dengan asumsi independensi antar fitur. Algoritma ini dikenal karena kesederhanaannya dan kemampuannya menangani dataset besar dengan cepat. Meskipun sering kali asumsi independensi tidak berlaku dalam dunia nyata, Naïve Bayes tetap efektif untuk banyak aplikasi klasifikasi teks, termasuk analisis sentimen. Gaussian Naïve Bayes digunakan untuk data numerik kontinu, sementara Multinomial Naïve Bayes lebih cocok untuk data diskrit seperti hitungan kata dalam teks [18].

b. *Logistic Regression*

Logistic Regression adalah metode regresi yang digunakan untuk klasifikasi, baik biner maupun multi-kelas. Model ini menggunakan fungsi logistik untuk menghasilkan probabilitas antara 0 dan 1, yang kemudian dipakai untuk klasifikasi. Berbeda dengan Naïve Bayes, yang mengasumsikan independensi antar fitur, Logistic Regression lebih fleksibel dalam memodelkan hubungan antar fitur. Karena itu, Logistic Regression lebih cocok untuk data yang lebih kompleks, di mana fitur-fitur saling berkorelasi dengan kuat.

Beberapa penelitian menunjukkan bahwa Logistic Regression sering kali lebih unggul dibandingkan Naïve Bayes pada dataset yang lebih kompleks, di mana hubungan antar fitur lebih rumit dan saling berkorelasi.

3. Metode yang Diusulkan

Metode penelitian ini menjelaskan langkah-langkah yang dilakukan dalam menerapkan algoritma Naïve Bayes dan Logistic Regression untuk analisis sentimen pada ulasan sebuah toko produk skincare di Tokopedia. Dalam bagian ini, akan dijelaskan tentang pendekatan yang digunakan, pengumpulan data, serta tahapan yang terlibat dalam pemrosesan dan analisis data.

3.1 Pendekatan Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan tujuan untuk melakukan sebuah analisis sentimen terhadap ulasan produk skincare yang terdapat di dalam platform e-commerce Tokopedia. Pendekatan yang digunakan dalam pengaplikasian teknik Natural Language Processing (NLP) dan Machine Learning untuk mengaplikasikan ulasan menjadi kategori sentimen positif, negative, netral. Dua algoritma yang digunakan Adalah Naïve Bayes dan juga Logistic Regression, yang nantinya akan dibandingkan kinerjanya untuk mengklasifikasikan sentimen berdasarkan dataset ulasan

3.2 Identifikasi Masalah

Permasalahan utama dalam penelitian ini Adalah bagaimana cara melakukan analisis terhadap toko skincare pada platform Tokopedia. Ulasan ini mengandung opini yang beragam dan perlu dianalisis untuk membantu penjual dan pembeli memahami sentimen yang ada di pasar. Penelitian ini bertujuan untuk mengidentifikasi dan mengklasifikasikan sentimen ulasan secara otomatis menggunakan dua metode pembelajaran mesin yang telah terbukti efektif.

3.3 Pengumpulan Data

Data dalam penelitian ini didapatkan menggunakan teknik web scraping pada sebuah toko skincare di platform e-commerce Tokopedia, menggunakan Pustaka Python BeautifulSoup. Teknik ini memungkinkan untuk mengumpulkan ulasan produk secara otomatis, yang selanjutnya digunakan untuk analisis sentimen. Proses pengumpulan data ini memastikan bahwa informasi yang diperoleh bersifat terkini dan relevan. Sebagai referensi, Qadri et al (2024) menyatakan bahwa teknik scraping merupakan metode yang efektif untuk memperoleh data ulasan dari platform e-commerce seperti Shopee, yang dapat diolah lebih lanjut menggunakan algoritma machine learning untuk klasifikasi sentimen [19].

3.4 Ekplorasi dan Pra-Pemrosesan Data (EDA)

Tahap ini bertujuan untuk membersihkan dan mempersiapkan data ulasan mentah agar dapat diolah secara efektif oleh model machine learning. Proses ini sangat krusial karena kualitas data input akan sangat memengaruhi kinerja model klasifikasi. Berdasarkan naskah Anda, langkah-langkah yang dilakukan adalah sebagai berikut:

- a. **Pembersihan Teks:** Proses ini mencakup beberapa sub-tahapan yang esensial contohnya adalah:
 - 1) Case Folding: Seluruh teks ulasan diubah menjadi format huruf kecil (lowercase) untuk memastikan konsistensi dan menghindari duplikasi fitur akibat perbedaan kapitalisasi (misalnya, "Produk" dan "produk" dianggap sebagai kata yang sama).
 - 2) Penghapusan Angka dan Tanda Baca: Semua karakter numerik dan tanda baca (seperti titik, koma, tanda seru) dihilangkan dari teks. Langkah ini bertujuan untuk mengurangi noise dan memfokuskan analisis pada kata-kata yang membawa sentimen.
- b. **Tokenisasi:** Setelah teks bersih, setiap kalimat ulasan dipecah menjadi unit-unit kata individual yang disebut token. Tokenisasi merupakan langkah fundamental dalam analisis teks untuk memungkinkan pemrosesan lebih lanjut pada tingkat kata.
 - 1) Stopwords Removal: Kata-kata umum yang sering muncul namun tidak memiliki makna sentimen yang signifikan (misalnya, "yang", "dan", "di", "itu") dihapus dari daftar token. Proses ini membantu mengurangi dimensi fitur dan memfokuskan model pada kata-kata yang lebih informatif.
 - 2) Stemming/Lemmatization: Kata-kata yang telah ditokenisasi diubah ke bentuk dasarnya untuk mengurangi variasi morfologis. Meskipun disebutkan, implementasi pada model yang Anda sediakan tampaknya tidak menyertakan langkah ini untuk menjaga makna asli dari setiap kata dalam konteks ulasan skincare.

3.5 Ekstraksi Fitur: TF-IDF

Setelah data teks melalui tahap pra-pemrosesan, langkah selanjutnya adalah mengubah teks menjadi representasi numerik yang dapat dipahami oleh algoritma machine learning. Dalam penelitian ini, metode yang digunakan adalah Term Frequency-Inverse Document Frequency (TF-IDF).

TF-IDF adalah sebuah metode statistik yang digunakan untuk mengukur seberapa penting sebuah kata dalam suatu dokumen di dalam sebuah koleksi atau korpus. Bobot TF-IDF meningkat seiring dengan frekuensi kemunculan sebuah kata dalam dokumen (TF), namun diimbangi dengan frekuensi kemunculan kata tersebut di seluruh korpus dokumen (IDF). Hal ini memastikan bahwa kata-kata yang unik dan signifikan untuk sebuah ulasan akan memiliki bobot yang lebih tinggi. Pada implementasinya, digunakan TfidfVectorizer dari pustaka Scikit-learn untuk mengubah korpus ulasan yang telah dibersihkan menjadi matriks fitur TF-IDF. Data kemudian dibagi menjadi data latih (training set) dan data uji (testing set) dengan proporsi 80:20 untuk melatih dan mengevaluasi model.

3.6 Pemodelan dan Pelatihan Algoritma

Penelitian ini membandingkan kinerja dua algoritma klasifikasi yang umum digunakan dalam analisis sentimen.

3.6.1 Algoritma Naïve Bayes

Model pertama yang diimplementasikan adalah Naïve Bayes Classifier. Algoritma ini merupakan metode klasifikasi probabilistik yang didasarkan pada Teorema Bayes dengan asumsi independensi antar fitur. Dalam konteks analisis teks, model ini menghitung probabilitas sebuah ulasan masuk ke dalam kategori sentimen (positif atau negatif) berdasarkan kata-kata yang ada di dalamnya. Variasi yang digunakan adalah MultinomialNB, yang sangat cocok untuk data diskrit seperti frekuensi kemunculan kata dalam teks.

3.6.2 Algoritma Logistic Regression

Model kedua adalah Logistic Regression, sebuah metode statistik yang digunakan untuk klasifikasi biner. Model ini memprediksi probabilitas sebuah ulasan diklasifikasikan sebagai sentimen positif atau negatif dengan menggunakan fungsi logistik. Berbeda dari Naïve Bayes, Logistic Regression mampu memodelkan hubungan yang lebih kompleks antar fitur. Untuk mendapatkan performa optimal, model Logistic Regression dilatih menggunakan GridSearchCV, sebuah teknik untuk melakukan hyperparameter tuning secara sistematis. Parameter yang dioptimalkan mencakup max_features dan ngram_range untuk TF-IDF, serta parameter regularisasi C dan class_weight untuk model Logistic Regression itu sendiri.

3.7 Metrix Evaluasi Kinerja

Untuk mengukur dan membandingkan kinerja kedua model secara objektif, penelitian ini menggunakan beberapa metrik evaluasi standar, yaitu:

- a. Accuracy: Mengukur persentase ulasan yang berhasil diklasifikasikan dengan benar oleh model dari total keseluruhan data uji.
- b. Confusion Matrix: Sebuah tabel yang merangkum hasil prediksi dengan membandingkannya terhadap kelas aktual. Matriks ini menyajikan jumlah True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN).
- c. Precision, Recall, dan F1-Score:
 - 1) Precision: Mengukur tingkat ketepatan dari ulasan yang diprediksi sebagai positif.
 - 2) Recall (Sensitivity): Mengukur kemampuan model untuk menemukan kembali semua ulasan yang sebenarnya positif.
 - 3) F1-Score: Rata-rata harmonik dari Precision dan Recall, memberikan skor tunggal yang menyeimbangkan kedua metrik tersebut.
- d. ROC-AUC: Kurva Receiver Operating Characteristic (ROC) memvisualisasikan performa model dalam membedakan kelas positif dan negatif pada berbagai ambang batas. Area Under the Curve (AUC) merepresentasikan skor agregat dari kinerja tersebut; semakin mendekati 1, semakin baik kemampuan model dalam melakukan klasifikasi.

4. Hasil dan Pembahasan

Bab ini menyajikan hasil dari analisis sentimen ulasan produk skincare di Tokopedia dengan menggunakan dua model klasifikasi, yaitu Naïve Bayes dan Regresi Logistik. Pembahasan mencakup hasil dari setiap tahapan, mulai dari pra-pemrosesan data, hasil pemodelan dan evaluasi masing-masing algoritma, hingga perbandingan kinerja keduanya untuk menentukan model yang paling efektif.

4.1 Eksplorasi dan Pra Pemrosesan Data

Data ulasan yang diperoleh melalui scraping melalui serangkaian tahap pra-pemrosesan untuk membersihkan dan menstandarisasi data teks. Tahapan ini, yang meliputi Case Folding, penghapusan angka/tanda baca, Tokenizing, dan Stopwords Removal, sangat krusial untuk memastikan kualitas data input. Hasil visualisasi dari tahapan tersebut ditunjukkan pada Tabel 1.

Tabel 1. Ini adalah sebuah tabel. Tabel harus ditempatkan dalam teks utama, dekat dengan tempat pertama kali tabel tersebut dirujuk.

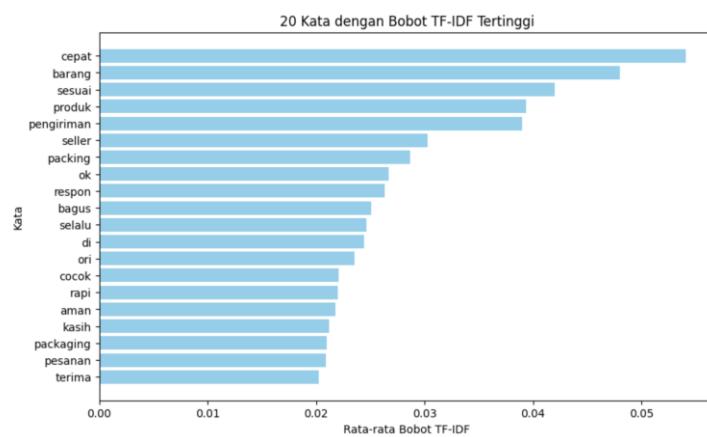
Tahap	Contoh Teks Hasil
Teks Awal (Ulasan)	Produk baik\nPacking baik\nPengiriman baik
Case Folding	produk baik\npacking baik\\pengiriman baik
Pembersihan Angka & Tanda Baca	produk baik packing baik pengiriman baik
Tokenizing	['produk', 'baik', 'packing', 'baik', 'pengiriman', 'baik']
Stopwords Removal	['produk', 'baik', 'packing', 'baik', 'pengiriman', 'baik']

Teks Bersih Final produk baik packing baik pengiriman baik

Setelah proses pembersihan, dataset final yang digunakan untuk pemodelan terdiri dari 973 ulasan yang telah bersih dan siap diolah.

4.2 Ekstraksi Fitur IDF

Setelah data bersih, dilakukan ekstraksi fitur menggunakan Term Frequency – Inverse Document Frequency (TF-IDF). Metode ini mengubah teks ke dalam bentuk numerik sehingga dapat diproses oleh algoritma machine learning. Dataset dibagi menjadi data latih (778 sampel, 80%) dan data uji (195 sampel, 20%). Hasil representasi menghasilkan 1.102 fitur unik pada matriks TF-IDF. Untuk mengetahui kata yang paling dominan, dibuat visualisasi 20 kata dengan bobot TF-IDF tertinggi. Kata-kata seperti “produk”, “baik”, “pengiriman”, dan “ori” mendominasi ulasan positif, sementara kata “buruk” dan “kecewa” sering muncul pada ulasan negatif.

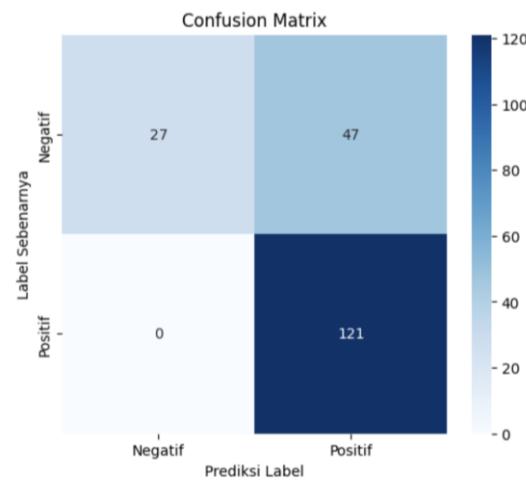


Gambar 1. 20 Kata Dengan Bobot TF-IDF Tertinggi

4.3 Pemodelan Algoritma

4.3.1 Algoritma Naïve Bayes

Algoritma Multinomial Naïve Bayes diimplementasikan sebagai model pertama dan menghasilkan akurasi sebesar 76% pada data uji. Hasil evaluasi model ini menunjukkan kinerja yang tidak seimbang, ditunjukkan melalui Confusion Matrix pada Gambar 2.



Gambar 2. Confusion Matrix Naïve Bayes

Seperti terlihat pada Gambar 2, kinerja model ini menunjukkan bias kuat ke kelas positif. Dari 74 ulasan negatif aktual, model salah mengklasifikasikan 47 ulasan sebagai positif (False Positive). Sebaliknya, model berhasil memprediksi sebagian besar ulasan positif dengan benar. Bias ini dikonfirmasi oleh Tabel 2, di mana recall kelas negatif hanya 0.36, berbanding terbalik

dengan recall kelas positif yang mencapai 1.00. Ini menunjukkan NB kurang andal dalam mengenali ulasan negatif.

Tabel 2. Laporan Klasifikasi Naïve Bayes

Kelas	Precision	Recall	F1-score	Support
Negatif	1.00	0.36	0.53	74
Positif	0.72	1.00	0.84	121
Accuracy		0.76		195

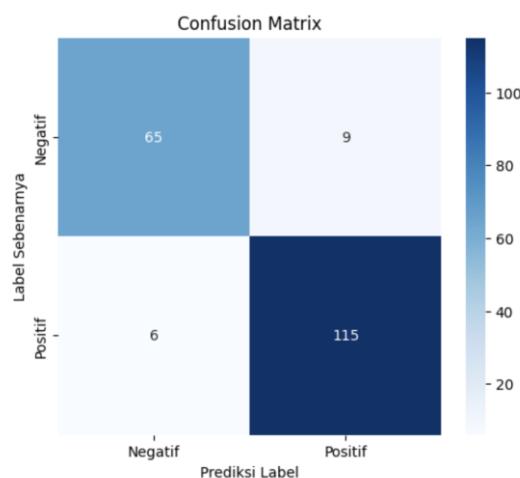
Dari Tabel dapat dilihat bahwa algoritma Naïve Bayes memperoleh akurasi 76%. Nilai precision untuk kelas negatif mencapai 1.00, artinya semua ulasan yang diprediksi negatif benar-benar negatif. Namun, nilai recall kelas negatif hanya 0.36, menunjukkan banyak ulasan negatif yang gagal dikenali (false negative cukup tinggi). Sementara itu, untuk kelas positif, model memperoleh precision 0.72 dan recall 1.00, yang berarti hampir semua ulasan positif berhasil terdeteksi. Secara keseluruhan, nilai f1-score kelas positif (0.84) jauh lebih tinggi dibanding kelas negatif (0.53). Temuan ini memperkuat hasil confusion matrix sebelumnya, yaitu bahwa Naïve Bayes cenderung lebih baik dalam mengklasifikasikan ulasan positif dibanding ulasan negatif.

4.3.2 Algoritma Logistic Regression

Model kedua menggunakan Logistic Regression dengan optimasi hyperparameter melalui GridSearchCV. Parameter terbaik yang diperoleh adalah:

- C = 10.0
- class_weight = balanced
- max_features = 1000
- ngram_range = (1,2)

Dengan parameter tersebut, Logistic Regression menghasilkan performa yang lebih unggul dibanding Naïve Bayes, dengan akurasi mencapai 92% pada data uji.



Gambar 3. Confusion Matrix Logistic Regression

Berdasarkan Gambar 3, dapat dilihat bahwa model Logistic Regression jauh lebih seimbang dalam mengenali ulasan positif maupun negatif. Dari 74 data ulasan negatif, sebanyak 65 ulasan diprediksi benar sebagai negatif, sedangkan hanya 9 ulasan salah diklasifikasikan sebagai positif. Untuk 121 ulasan positif, model memprediksi dengan benar sebanyak 115 ulasan, sementara 6 ulasan salah terkласifikasi sebagai negatif. Hal ini menunjukkan bahwa

Logistic Regression tidak hanya mampu mengenali ulasan positif dengan baik, tetapi juga cukup akurat dalam mengidentifikasi ulasan negatif.

Tabel 3. Laporan Klasifikasi Logistic Regression

Kelas	Precision	Recall	F1-score	Support
Negatif	0.92	0.88	0.90	74
Positif	0.93	0.95	0.94	121
Accuracy	0.92		195	

Tabel 3 memperlihatkan detail evaluasi model Logistic Regression. Untuk kelas negatif, diperoleh precision 0.92, recall 0.88, dan f1-score 0.90. Artinya, sebagian besar ulasan negatif berhasil diprediksi dengan benar, walaupun masih ada sebagian kecil yang salah klasifikasi. Sementara itu, untuk kelas positif, model mencapai precision 0.93, recall 0.95, dan f1-score 0.94, yang menandakan kemampuan model sangat baik dalam mendeteksi ulasan positif dengan tingkat kesalahan yang sangat kecil. Secara keseluruhan, Logistic Regression memperoleh akurasi total 92% dan nilai Macro Average F1-score 0.92. Nilai yang tinggi dan seimbang ini menegaskan bahwa model mampu mengklasifikasikan kedua kelas (positif dan negatif) secara merata, menjadikannya jauh lebih stabil dibandingkan Naïve Bayes.

4.4 Matrix Evaluasi Kinerja Model

Untuk menilai performa kedua algoritma, dilakukan perbandingan berdasarkan metrik akurasi, precision, recall, dan f1-score.

Tabel 4. Perbandingan Kinerja Model

Algoritma	Akurasi	Precision	Recall	F1-score
Naïve Bayes	0.759	0.72	0.68	0.72
Logistic Regression	0.923	0.93	0.91	0.92

Dari tabel di atas, terlihat bahwa Logistic Regression secara konsisten mengungguli Naïve Bayes pada semua metrik. Logistic Regression lebih mampu menangkap hubungan antar fitur dan bekerja baik pada dataset dengan korelasi antar kata.

5. Perbandingan

Hasil analisis menunjukkan perbedaan kinerja yang signifikan antara algoritma Naïve Bayes (NB) dan Logistic Regression (LR) dalam mengklasifikasikan sentimen ulasan produk skincare di Tokopedia. Tabel 4 memperjelas bahwa LR secara konsisten mengungguli NB pada semua metrik evaluasi.

Perbedaan utama kinerja ini terletak pada asumsi dasar kedua model. Naïve Bayes, yang mengasumsikan independensi fitur (kata), menunjukkan bias yang kuat ke kelas positif pada dataset ini. Hal ini terlihat dari recall kelas negatif yang sangat rendah (0.36) dan nilai Macro Average F1-score NB yang hanya sekitar 0.69. Kinerja yang tidak seimbang ini mengindikasikan bahwa asumsi independensi fitur Naïve Bayes tidak terpenuhi atau terpengaruh oleh ketidakseimbangan kelas minoritas (negatif) dalam data.

Sebaliknya, Logistic Regression terbukti lebih adaptif dan stabil. Dengan menggunakan Macro Average F1-score 0.92, LR berhasil mencapai klasifikasi yang seimbang antar kelas (F1-score negatif 0.90 dan positif 0.94). Keunggulan LR ini disebabkan oleh kemampuannya memodelkan hubungan antar fitur yang saling berkorelasi, menjadikannya lebih tepat untuk analisis sentimen pada dataset ulasan e-commerce yang kompleks.

Dengan demikian, meskipun NB unggul pada kecepatan dan kesederhanaan implementasi, LR lebih tepat digunakan pada data heterogen dengan tingkat kompleksitas dan korelasi fitur yang tinggi, seperti ulasan produk skincare ini.

6. Kesimpulan

Berdasarkan hasil analisis perbandingan, penelitian ini menyimpulkan bahwa Logistic Regression (LR) lebih unggul dibandingkan Naïve Bayes (NB) dalam menganalisis sentimen ulasan produk skincare di Tokopedia. LR mampu menangani data dengan korelasi antar fitur yang kompleks, sehingga lebih stabil untuk dataset heterogen dengan keragaman bahasa, sedangkan NB tetap relevan karena sederhana, cepat, dan akurat pada dataset besar dengan kebutuhan klasifikasi instan. Oleh karena itu, LR direkomendasikan sebagai metode utama, sementara NB dapat menjadi alternatif sesuai kebutuhan. Untuk pengembangan ke depan, penelitian dapat diarahkan pada model hybrid yang mengombinasikan LR dengan algoritma machine learning lain serta mempertimbangkan variabel eksternal seperti aspek demografis dan perilaku konsumen guna meningkatkan akurasi analisis sentimen.

Ucapan Terimakasih

Puji syukur saya panjatkan kehadirat Tuhan Yang Maha Esa atas rahmat dan karunia-Nya, saya dapat menuntaskan penelitian ini hingga tersusun menjadi naskah ilmiah. Saya berhutang banyak rasa terima kasih kepada Ibu Indah Triaty dan Ayah Ifan Setiadi yang selalu memberikan doa, perhatian, dan dukungan tanpa henti sepanjang perjalanan ini. Saya juga menyampaikan penghargaan mendalam kepada Bapak Yunus Widjaja, S.Kom., M.M., yang telah membimbing saya dengan arahan yang jelas, saran yang berharga, serta kesabaran dalam setiap proses penyusunan penelitian

Tidak lupa, saya berterima kasih kepada sahabat-sahabat saya yang selalu ada dan membantu, baik dalam memberikan semangat, berdiskusi, maupun mendampingi ketika menghadapi kesulitan. Akhir kata, segala bantuan, doa, dan dorongan dari semua pihak yang tidak dapat saya sebutkan satu per satu, menjadi bagian penting yang menguatkan saya hingga penelitian ini dapat terselesaikan dengan baik.

Daftar Pustaka

- [1] T. Mustajibah and A. Trilaksana, “Dinamika E-Commerce di Indonesia Tahun 1999-2015,” AVATARA, e-Journal Pendidikan Sejarah, vol. 10, no. 3, 2021.
- [2] S. M. Salsabila, A. A. Murtopo, and N. Fadhilah, “Analisis Sentimen Pelanggan Tokopedia Menggunakan Metode Naïve Bayes Classifier,” Jurnal Minfo Polgan, vol. 11, no. 02, pp. 30–35, 2022.
- [3] M. Aulia and A. Hermawan, “Analisis Perbandingan Algoritma SVM, Naïve Bayes, dan Perceptron untuk Analisis Sentimen Ulasan Produk Tokopedia,” Jurnal Media Informatika Budidarma, vol. 7, no. 4, pp. 1850–1859, 2023.
- [4] B. Z. Ramadhan, I. Riza, and I. Maulana, “Analisis Sentimen Ulasan Pada Aplikasi E-Commerce Dengan Menggunakan Algoritma Naïve Bayes,” Journal of Applied Informatics and Computing (JAIC), vol. 6, no. 2, pp. 220–225, 2022.
- [5] A. Muzaki, V. Febriana, and W. N. Cholifah, “Analisis Sentimen pada Ulasan Produk di E-commerce dengan Metode Naïve Bayes,” Jurnal Riset dan Aplikasi Mahasiswa Informatika (JRAMI), vol. 05, no. 04, pp. 758–765, 2024.
- [6] R. Apriani and D. Gustian, “Analisis Sentimen dengan Naïve Bayes Terhadap Komentar Aplikasi Tokopedia,” Jurnal Rekayasa Teknologi Nusa Putra, vol. 6, no. 1, pp. 54–62, 2019.
- [7] D. Nasien et al., “Perbandingan Implementasi Machine Learning Menggunakan Metode KNN, Naïve Bayes, Dan Logistik Regression Untuk Mengklasifikasi Penyakit Diabetes,” Jurnal Teknik Informatika, vol. 4, no. 1, 2024.
- [8] O. N. Cahyani and F. Budiman, “Performa Logistic Regression dan Naïve Bayes dalam Klasifikasi Berita Hoax di Indonesia,” Edumatic: Jurnal Pendidikan Informatika, vol. 9, no. 1, pp. 60–68, 2025.
- [9] A. G. Wastradmojo and R. Kusumastuti, “Analisa Sentimen Ulasan di Tokopedia dengan Metode Naïve Bayes,” in Seminar Nasional Amikom Surakarta (SEMNASA), 2024, pp. 369–382.
- [10] D. Fadila and M. Ikhsan, “Analisis Sentimen Pada Aplikasi Tokopedia Menggunakan Metode Support Vector Machine,” Progresif: Jurnal Ilmiah Komputer, vol. 21, no. 1, pp. 238–247, 2025.
- [11] Ayumi, A. Fayola, R. Darianty, and Stephanie, “Analisis Sentimen Ulasan Produk Kecantikan di Tokopedia Menggunakan In- doBERT,” ISAJ, vol. 01, no. 01, pp. 11–18, 2025.
- [12] A. A. Lestari, A. Faqih, and G. Dwilestari, “Improving Sentiment Analysis Performance of Tokopedia Reviews Using Principal Component Analysis and Naïve Bayes Algorithm,” Journal of Artificial Intelligence and Engineering Applications, vol. 4, no. 2, pp. 759-763, 2025.

- [13] E. I. Sitorus and M. Simanjuntak, "Pengaruh Review Online terhadap Keputusan Pembelian Produk Skincare pada Platform E-Commerce di Indonesia (Study Kasus: Mahasiswa Institut Teknologi Del)," *Jurnal Pendidikan Tambusai*, vol. 8, no. 3, pp. 50580–50589, 2024.
- [14] S. Zulaika, "Pengaruh Harga dan Rating Produk Terhadap Keputusan Pembelian Produk Berdasarkan Data Sekunder di E-Commerce," *KEIZAI: Jurnal Kajian Ekonomi, Manajemen, & Akuntansi*, vol. 5, no. 2, pp. 199–213, 2024.
- [15] E. Hasibuan and E. A. Heriyanto, "Analisis Sentimen pada Ulasan Aplikasi Amazon Shopping di Google Play Store Menggunakan Naive Bayes Classifier," *JTS Jurnal Teknik dan Science*, vol. 1, no. 3, pp. 13–24, 2022.
- [16] Ayumi, A. Fayola, R. Darianty, and Stephanie. "Analisis Sentimen Ulasan Produk Kecantikan di Tokopedia Menggunakan In- doBERT." *ISAJ*, vol. 01, no. 01, 2025, pp. 11-18
- [17] Mustajibah, T. and Trilaksana, A. "Dinamika E-Commerce di Indonesia Tahun 1999-2015." *AVATARA, e-Jurnal Pendidikan Sejarah*, vol. 10, no. 3, 2021
- [18] Bahtiar, S. A. H. "Comparison of Naïve Bayes and Logistic Regression in Sentiment Analysis on Marketplace Reviews Using Rating-Based Labeling." *Journal of Information Systems and Informatics*, vol. 4, no. 3, 2023
- [19] Qadri, N. A., Alam, S., & Taufik, M. (2024). Analisis Sentimen Brand Barenbliss Menggunakan Machine Learning (Studi Kasus: E-Commerce Shopee). *Innovative: Journal Of Social Science Research*, 5(1), 2057-2069.
- [20] W. Wahyuni, "Analisis Sentimen terhadap Opini Feminisme Menggunakan Metode Naïve Bayes," *Jurnal Informatika Ekonomi Bisnis*, vol. 4, no. 4, pp. 148–153, 2022, doi: 10.37034/infeb.v4i4.162.