

ANALISIS SENTIMEN ULASAN PRODUK E-COMMERCE MENGUNAKAN NAIVE BAYES

Leo Nakanisi Aran^{1*}, Andrian Anye², Sophia Daud³, Vinsensia Florince Seke⁴, Aldi Bastiatul Fawait⁵

¹²³⁴⁵ Program Studi Ilmu Komputer, Fakultas Teknik Informatika, Universitas Widya Gama Mahakam Samarinda, Indonesia

*Corresponding Author. Leonakanisarian@gmail.com

Receipt: Juni, 2026 ; Accepted:

Abstract The rapid growth of e-commerce in Indonesia has significantly increased the number of customer reviews on marketplace platforms, particularly Tokopedia. These reviews contain valuable customer opinions that can be utilized to evaluate product quality and service performance. However, the large volume of reviews makes manual analysis inefficient and time-consuming. This study aims to implement the **Multinomial Naive Bayes** algorithm for sentiment analysis of Tokopedia product reviews to automatically classify customer opinions into positive and negative sentiments. The dataset used consists of **500 labeled customer reviews** obtained from Kaggle. The research process includes text preprocessing through **case folding** and **cleaning**, feature extraction using **Term Frequency-Inverse Document Frequency (TF-IDF)** with **max_features=3000** and **ngram_range=(1,2)**, followed by data splitting into **80% training data** and **20% testing data**. The classification process is performed using the **Multinomial Naive Bayes** algorithm. Model performance is evaluated using **Accuracy, Precision, Recall, F1-Score, and Confusion Matrix**. The experimental results achieved an **accuracy of 83%**, indicating that the combination of preprocessing techniques, TF-IDF feature extraction, and the Multinomial Naive Bayes algorithm is effective in classifying customer sentiments on Tokopedia product reviews and can support the evaluation of product quality and service performance in e-commerce platforms.

Abstrak: Perkembangan e-commerce di Indonesia menyebabkan jumlah ulasan pelanggan pada platform marketplace, khususnya Tokopedia, terus meningkat. Ulasan tersebut memuat opini pelanggan yang dapat dimanfaatkan untuk mengevaluasi kualitas produk dan layanan. Namun, jumlah ulasan yang besar menyebabkan proses analisis secara manual menjadi kurang efektif dan memerlukan waktu yang cukup lama. Penelitian ini bertujuan untuk menerapkan algoritma Multinomial Naive Bayes dalam melakukan analisis sentimen terhadap ulasan produk Tokopedia sehingga ulasan dapat diklasifikasikan ke dalam sentimen positif dan negatif secara otomatis. Dataset yang digunakan berasal dari Kaggle sebanyak 500 data ulasan yang telah diberi label sentimen. Tahapan penelitian meliputi preprocessing data melalui case folding dan cleaning, ekstraksi fitur menggunakan Term Frequency-Inverse Document Frequency (TF-IDF) dengan parameter max_features=3000 dan ngram_range=(1,2), pembagian data menggunakan rasio 80% data training dan 20% data testing, kemudian dilakukan proses klasifikasi menggunakan algoritma Multinomial Naive Bayes. Performa model dievaluasi menggunakan Accuracy Score, Precision, Recall, F1-Score,

dan Confusion Matrix. Hasil penelitian menunjukkan bahwa model memperoleh akurasi sebesar 83%, sehingga menunjukkan bahwa kombinasi preprocessing, TF-IDF, dan algoritma Multinomial Naive Bayes mampu mengklasifikasikan sentimen ulasan pelanggan Tokopedia dengan baik serta dapat digunakan sebagai pendukung evaluasi kualitas produk dan layanan pada platform e-commerce.

Keywords: Analisis Sentimen, Tokopedia, *Machine Learning*, *Multinomial Naive Bayes*, *TF-IDF*, *E-Commerce*.

How to Cite this Article (IEEE):

S. A. Alfianto, Yudiansyah, and Aldi, "Article Title Here," *Journal of Multidisciplinary Informatics, Artificial Intelligence, and Business Innovation.*, vol. 1, no. 21, p. 1-5, 2026, doi: 10.24903/minabis.v1i1.xxx

PENDAHULUAN

Perkembangan dan pertumbuhan e-commerce di Indonesia salah satunya termasuk Tokopedia menjadi salah satu sarana utama masyarakat dalam melakukan transaksi jual beli secara daring[1]. Selain menyediakan berbagai pilihan produk, Tokopedia juga menyediakan fitur ulasan pelanggan yang memungkinkan konsumen memberikan penilaian dan pengalaman setelah melakukan pembelian[2]. Ulasan tersebut menjadi sumber informasi penting bagi calon pembeli maupun penjual dalam mengevaluasi kualitas produk dan layanan yang tersedia.

Salah satu pendekatan populer bagi masyarakat Indonesia adalah analisis sentimen yang dapat membantu mengklasifikasikan opini pengguna menjadi kategori positif dan negatif[3]. Salah satu metode yang sering ditemukan adalah Naive Bayes, metode ini digunakan karena kemampuannya yang sederhana, mudah dan efisien untuk memproses data menjadi sangat akurat[4]. Metode ini juga telah banyak digunakan dalam beberapa studi terdahulu dengan berbagai hasil yang memuaskan.

Ulasan pelanggan merupakan salah satu komponen penting dalam ekosistem e-commerce [5] berfungsi sebagai sarana komunikasi antara konsumen dan penjual serta menjadi referensi bagi calon pembeli dalam mengambil keputusan pembelian[6]. Selain itu, ulasan pelanggan dapat digunakan oleh penjual untuk mengevaluasi kualitas produk maupun layanan yang diberikan kepada konsumen. [7]

[7] Penelitian yang dilakukan pada ulasan pelanggan Tokopedia menggunakan algoritma Naive Bayes menunjukkan bahwa metode tersebut mampu mencapai akurasi sebesar 89% dalam mengklasifikasikan sentimen positif dan negatif. Selain itu, [7] pada ulasan pengguna Tokopedia memperoleh akurasi sebesar 94,5%, yang menunjukkan bahwa metode machine learning memiliki kemampuan yang baik dalam melakukan klasifikasi sentimen pada data teks.

Meskipun berbagai penelitian telah dilakukan, sebagian besar penelitian masih menggunakan dataset yang terbatas pada jenis produk tertentu atau jumlah data yang relatif kecil. Oleh karena itu, penelitian ini berfokus pada analisis sentimen ulasan produk pada platform Tokopedia menggunakan algoritma Naive Bayes dengan memanfaatkan data ulasan produk yang lebih spesifik sehingga dapat memberikan gambaran yang lebih mendalam mengenai persepsi pelanggan terhadap produk yang diteliti[8].

Berdasarkan uraian tersebut, rumusan masalah dalam penelitian ini adalah bagaimana penerapan algoritma Naive Bayes dalam mengklasifikasikan sentimen ulasan produk pada Tokopedia serta bagaimana tingkat performa algoritma tersebut dalam mengidentifikasi sentimen pelanggan.

Tujuan penelitian ini adalah menerapkan algoritma Naive Bayes untuk melakukan analisis sentimen terhadap ulasan produk pada platform Tokopedia, mengidentifikasi kecenderungan sentimen pelanggan, serta mengevaluasi performa algoritma menggunakan metrik akurasi, precision, recall, dan F1-score.⁶⁵

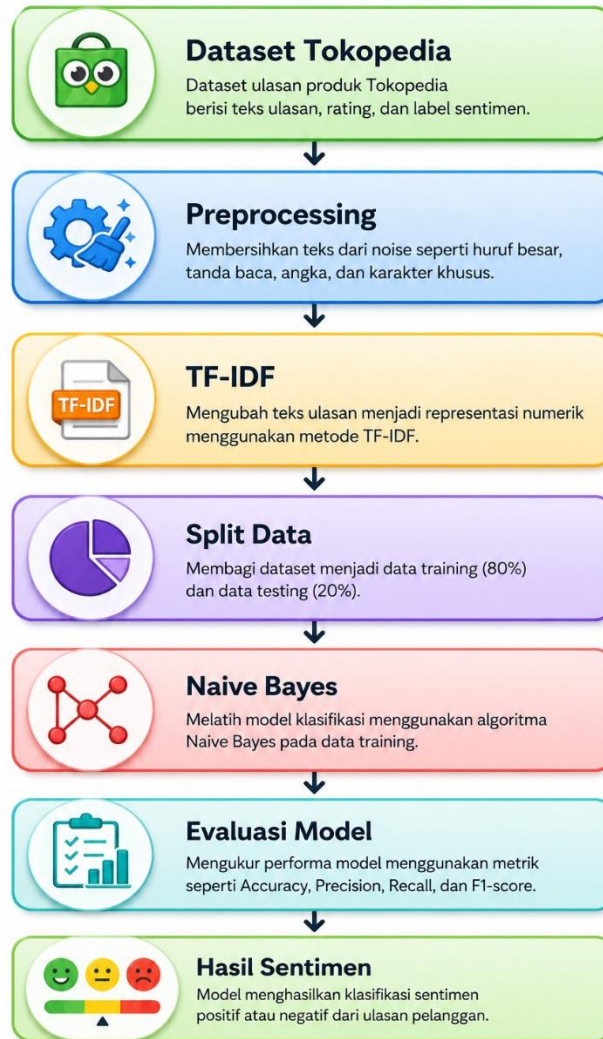
METODE

Penelitian ini menggunakan pendekatan kuantitatif dengan metode klasifikasi teks berbasis algoritma Naive Bayes. Tujuan penelitian adalah mengklasifikasikan sentimen ulasan pelanggan Tokopedia ke dalam dua kategori, yaitu sentimen positif dan sentimen negatif.

Dataset yang digunakan merupakan dataset ulasan pelanggan Tokopedia yang diperoleh dari platform Kaggle. Jumlah data yang digunakan dalam penelitian ini sebanyak 500 data ulasan yang terdiri dari teks ulasan pelanggan, rating, dan label sentimen. Label sentimen dibagi menjadi dua kelas, yaitu positif dan negatif.

Pengolahan dan analisis data dilakukan menggunakan bahasa pemrograman Python pada lingkungan Google Colab. Library yang digunakan meliputi pandas untuk membaca dan mengelola dataset, re (*Regular Expression*) untuk proses pembersihan teks, serta scikit-learn untuk proses ekstraksi fitur menggunakan TF-IDF, pembagian data training dan testing, penerapan algoritma Naive Bayes, dan evaluasi model^[9].

Tahapan preprocessing dilakukan melalui proses case folding dengan mengubah seluruh huruf menjadi huruf kecil serta cleaning untuk menghapus karakter yang tidak diperlukan seperti tanda baca, angka, dan simbol tertentu. Proses tersebut bertujuan untuk menghasilkan data teks yang lebih bersih sebelum dilakukan ekstraksi fitur menggunakan TF-IDF^[10]. Selanjutnya data diubah menjadi bentuk numerik menggunakan metode TF-IDF^[11]. Dataset kemudian dibagi menjadi data training dan data testing dengan perbandingan 80:20. Proses klasifikasi dilakukan menggunakan algoritma Multinomial Naive Bayes untuk menentukan sentimen positif dan negatif. Hasil klasifikasi dievaluasi menggunakan Accuracy Score, Classification Report, dan Confusion Matrix untuk mengetahui performa model yang dihasilkan.



Gambar 1. Alur Prosedur Penelitian Analisis Sentimen Ulasan Tokopedia Menggunakan Naive Bayes

Gambar ini menunjukkan tahapan penelitian yang dimulai dari pengumpulan dataset ulasan Tokopedia sebanyak 500 data, dilanjutkan dengan preprocessing, ekstraksi fitur menggunakan TF-IDF, pembagian data training dan testing, [7]

II.1 Machine Learning

Machine Learning merupakan salah satu cabang dari Artificial Intelligence (AI) yang memungkinkan sistem komputer untuk mempelajari pola dari data dan membuat prediksi atau keputusan tanpa harus diprogram secara eksplisit[12]. Machine Learning bekerja dengan memanfaatkan algoritma yang mampu mengenali hubungan antar data sehingga dapat digunakan untuk menyelesaikan berbagai permasalahan[13] seperti klasifikasi, prediksi, pengelompokan data, dan analisis sentimen[14]. Dalam bidang pengolahan bahasa alami (*Natural Language Processing*), [15]

Pada penelitian ini, Machine Learning digunakan untuk melakukan klasifikasi sentimen terhadap ulasan produk pada platform Tokopedia. Proses klasifikasi dilakukan dengan mempelajari pola kata yang terdapat pada data pelatihan sehingga model dapat

mengidentifikasi sentimen positif maupun negatif pada data ulasan yang belum pernah diproses sebelumnya[16].

II.2 Algoritma Naive Bayes

Naive Bayes merupakan salah satu algoritma klasifikasi yang berbasis probabilitas dan menggunakan Teorema Bayes sebagai dasar perhitungannya[17]. Algoritma ini memiliki asumsi bahwa setiap fitur bersifat independen terhadap fitur lainnya, sehingga proses perhitungan probabilitas menjadi lebih sederhana dan efisien[18]. Meskipun menggunakan asumsi independensi yang sederhana, Naive Bayes terbukti memiliki performa yang baik dalam berbagai kasus klasifikasi teks, termasuk analisis sentimen[2].

Keunggulan algoritma Naive Bayes terletak pada kemampuannya dalam menangani data berdimensi tinggi, proses pelatihan yang cepat, serta kebutuhan komputasi yang relatif rendah dibandingkan dengan beberapa algoritma klasifikasi lainnya. Oleh karena itu, algoritma ini banyak digunakan dalam penelitian analisis sentimen pada media sosial, ulasan produk, klasifikasi dokumen, serta berbagai permasalahan pengolahan teks lainnya.

Persamaan dasar Teorema Bayes ditunjukkan pada Persamaan :

$$P(C|X)=P(X)P(X|C)\times P(C)$$

Keterangan:

- $P(C | X)$ = probabilitas data X termasuk ke dalam kelas C
- $P(X | C)$ = probabilitas data X jika diketahui berada pada kelas C
- $P(C)$ = probabilitas awal kelas
- $P(X)$ = probabilitas data

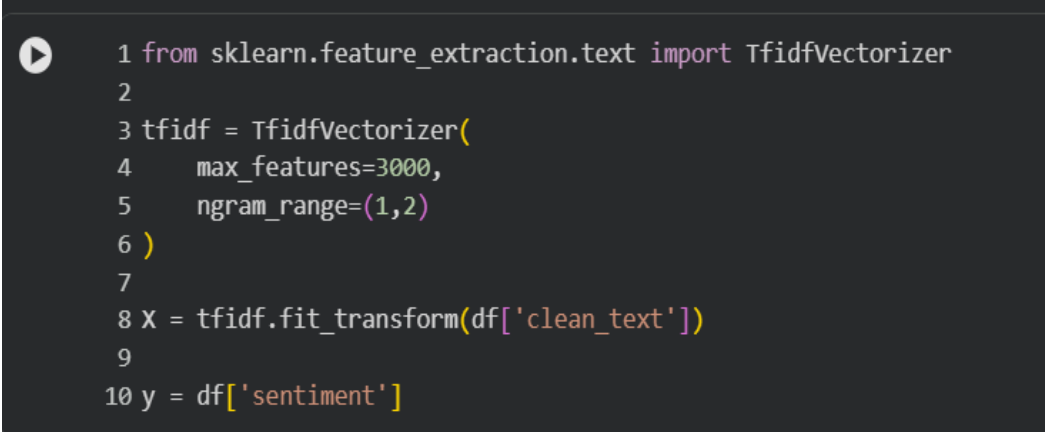
Pada penelitian ini digunakan metode **Multinomial Naive Bayes** karena metode tersebut sesuai untuk data teks yang telah direpresentasikan dalam bentuk bobot kata menggunakan TF-IDF. Algoritma ini digunakan untuk mengklasifikasikan ulasan pelanggan Tokopedia ke dalam dua kategori sentimen, yaitu sentimen positif dan sentimen negatif.

HASIL DAN PEMBAHASAN

3.1 Dataset Penelitian

Dataset yang digunakan dalam penelitian ini merupakan dataset ulasan pelanggan Tokopedia yang diperoleh dari platform Kaggle. Jumlah data yang digunakan sebanyak 500 data yang terdiri dari teks ulasan, rating, dan label sentimen. Label sentimen dibagi menjadi dua kategori, yaitu sentimen positif dan sentimen negatif.

Data tersebut digunakan untuk melatih dan menguji model Machine Learning menggunakan algoritma Naive Bayes. Sebelum dilakukan proses klasifikasi, data terlebih dahulu melalui tahap preprocessing dan ekstraksi fitur agar dapat diproses oleh model Machine Learning.



```

1 from sklearn.feature_extraction.text import TfidfVectorizer
2
3 tfidf = TfidfVectorizer(
4     max_features=3000,
5     ngram_range=(1,2)
6 )
7
8 X = tfidf.fit_transform(df['clean_text'])
9
10 y = df['sentiment']
    
```

Gambar 4. Kode tf-idf

Penggunaan parameter `max_features=3000` bertujuan untuk memilih 3000 kata yang paling penting sehingga model lebih fokus pada kata-kata yang memiliki pengaruh terhadap sentimen.

Sedangkan parameter `ngram_range=(1,2)` memungkinkan model untuk mengenali satu kata (unigram) maupun kombinasi dua kata (bigram).

Penggunaan bigram membuat model lebih memahami konteks suatu kalimat sehingga dapat meningkatkan performa klasifikasi.

3.4 Pembagian Data Training dan Testing

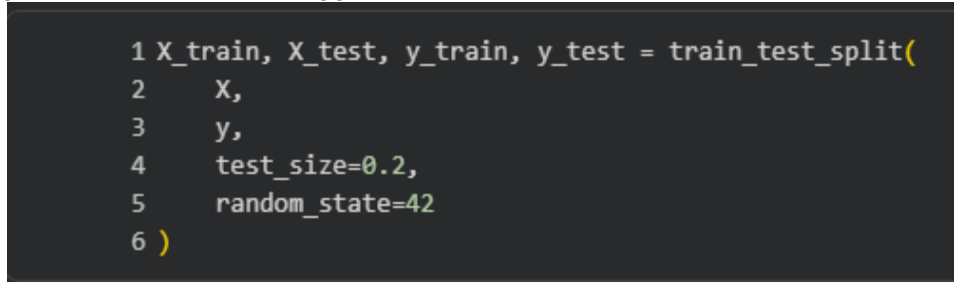
Dataset yang telah diubah ke bentuk numerik selanjutnya dibagi menjadi data training dan data testing menggunakan fungsi `train_test_split()`.

Pada penelitian ini digunakan pembagian data:

- 80% data training
- 20% data testing

Data training digunakan untuk melatih model Naive Bayes, sedangkan data testing digunakan untuk menguji kemampuan model dalam mengklasifikasikan sentimen yang belum pernah dipelajari sebelumnya.

Pembagian data dilakukan menggunakan kode:



```

1 X_train, X_test, y_train, y_test = train_test_split(
2     X,
3     y,
4     test_size=0.2,
5     random_state=42
6 )
    
```

Gambar 4. Kode train_test_split

Penggunaan `random_state=42` bertujuan agar pembagian data selalu konsisten sehingga hasil penelitian dapat direproduksi.

3.5 Hasil Klasifikasi Menggunakan Naive Bayes

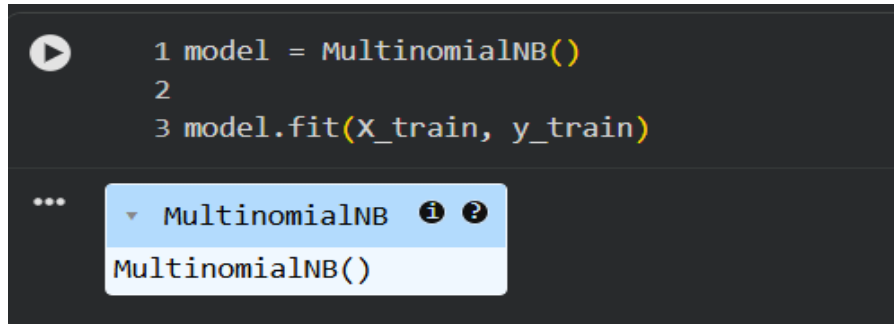
proses klasifikasi menggunakan algoritma **Multinomial Naive Bayes**.

Algoritma Naive Bayes dipilih karena memiliki proses perhitungan yang sederhana, cepat, dan efektif untuk klasifikasi data teks.

Pada penelitian ini, model mempelajari pola kata-kata yang muncul pada ulasan positif dan negatif. Setelah proses pelatihan selesai, model digunakan untuk memprediksi sentimen pada data testing.

Penggunaan algoritma Naive Bayes pada penelitian ini mampu menghasilkan performa yang cukup baik dalam mengelompokkan sentimen ulasan pelanggan Tokopedia.

Model dilatih menggunakan data training dengan kode:



```
1 model = MultinomialNB()  
2  
3 model.fit(X_train, y_train)
```

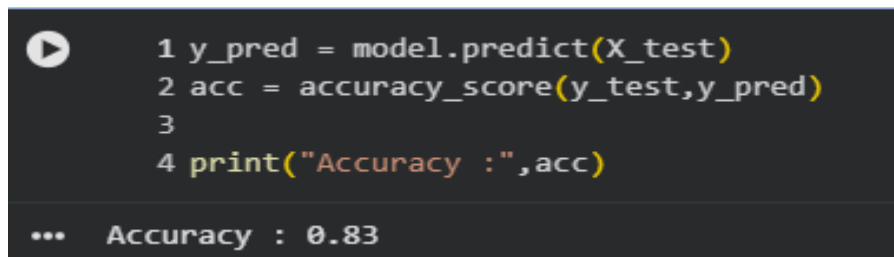
... MultinomialNB ⓘ ?
MultinomialNB()

Gambar 5. Kode multinomialNB()

3.6 Hasil Evaluasi Model

Evaluasi model dilakukan untuk mengetahui performa algoritma Naive Bayes dalam mengklasifikasikan sentimen ulasan pelanggan.

Berdasarkan hasil pengujian, model memperoleh nilai:



```
1 y_pred = model.predict(X_test)  
2 acc = accuracy_score(y_test, y_pred)  
3  
4 print("Accuracy :", acc)
```

... Accuracy : 0.83

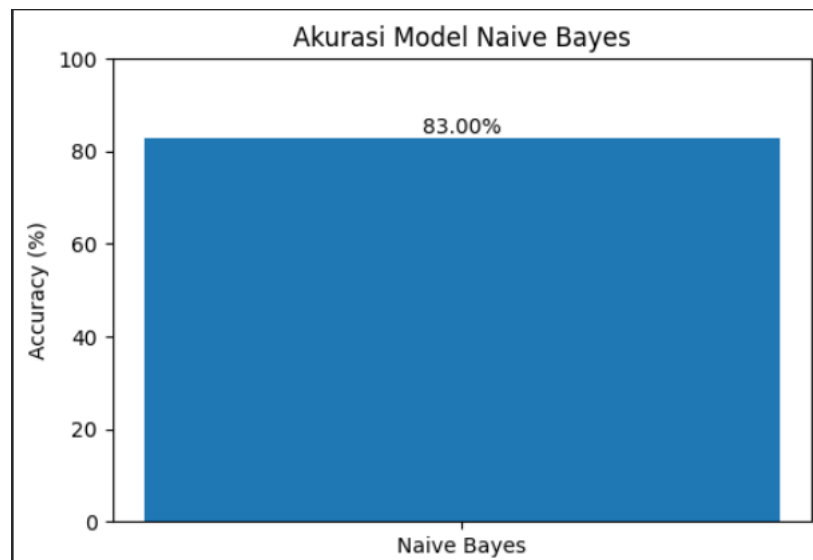
Gambar 6. Accuracy

Accuracy = 0.83 atau 83%

Evaluasi model dilakukan untuk mengetahui performa algoritma **Multinomial Naive Bayes** dalam mengklasifikasikan sentimen ulasan pelanggan Tokopedia. Evaluasi dilakukan menggunakan metrik **Accuracy**, **Precision**, **Recall**, **F1-Score**, dan **Confusion Matrix**.

Berdasarkan hasil pengujian, model memperoleh nilai **akurasi sebesar 0,83 atau 83%**. Hasil tersebut menunjukkan bahwa model mampu mengklasifikasikan sentimen ulasan pelanggan Tokopedia dengan tingkat ketepatan yang baik.

3.7 Hasil Evaluasi Model



Gambar 7. Grafik Accuracy

Berdasarkan Gambar Grafik Accuracy, model Naive Bayes memperoleh nilai akurasi sebesar 83%. Hasil tersebut menunjukkan bahwa model memiliki kemampuan yang baik dalam mengklasifikasikan sentimen ulasan pelanggan Tokopedia. Dengan tingkat akurasi sebesar 83%, sebagian besar data uji berhasil diprediksi dengan benar sesuai kategori sentimennya, yaitu positif dan negatif.

Dengan demikian, algoritma **Naive Bayes** yang dikombinasikan dengan metode **TF-IDF** dapat digunakan untuk melakukan analisis sentimen secara efektif pada dataset ulasan pelanggan Tokopedia yang digunakan dalam penelitian ini.

3.8 Pembahasan

Hasil penelitian menunjukkan bahwa algoritma Naive Bayes mampu melakukan klasifikasi sentimen ulasan pelanggan Tokopedia ke dalam kategori positif dan negatif dengan tingkat akurasi yang baik[20].

Keberhasilan model dipengaruhi oleh beberapa faktor, yaitu:

1. Kualitas preprocessing

Tahap preprocessing merupakan langkah awal yang sangat penting dalam proses analisis sentimen. Pada tahap ini, data ulasan yang masih berupa teks mentah dibersihkan dan disiapkan agar dapat diproses oleh algoritma machine learning. Proses preprocessing yang dilakukan meliputi pembersihan karakter khusus, penghapusan tanda baca, penghapusan angka yang tidak relevan, mengubah seluruh huruf menjadi huruf kecil (*case folding*), tokenisasi, penghapusan *stopwords*, serta stemming atau normalisasi kata.

Penerapan preprocessing memberikan pengaruh yang signifikan terhadap kualitas data karena mampu mengurangi noise yang terdapat dalam ulasan pengguna. Sebagai contoh, kata "Bagus!!!", "bagus", dan "BAGUS" akan diubah menjadi bentuk yang sama sehingga tidak dianggap sebagai kata yang berbeda oleh sistem. Dengan demikian, data menjadi lebih terstruktur dan konsisten.

Selain itu, preprocessing membantu mengurangi jumlah fitur yang tidak penting sehingga proses pelatihan model menjadi lebih efisien. Semakin baik kualitas preprocessing yang dilakukan, semakin baik pula kualitas informasi yang dapat

dipelajari oleh model klasifikasi. Oleh karena itu, preprocessing menjadi fondasi utama dalam membangun sistem analisis sentimen yang akurat.

2. Penggunaan metode TF-IDF

Setelah data melalui tahap preprocessing, langkah berikutnya adalah mengubah data teks menjadi bentuk numerik menggunakan metode **Term Frequency-Inverse Document Frequency (TF-IDF)**. Metode ini digunakan karena algoritma machine learning tidak dapat memproses data teks secara langsung.

TF-IDF bekerja dengan memberikan bobot pada setiap kata berdasarkan frekuensi kemunculannya dalam suatu dokumen dan tingkat keunikannya di seluruh dokumen. Kata yang sering muncul dalam satu ulasan tetapi jarang muncul pada ulasan lainnya akan mendapatkan bobot yang lebih tinggi. Sebaliknya, kata yang terlalu sering muncul pada seluruh dokumen akan memperoleh bobot yang lebih rendah.

Penggunaan TF-IDF memberikan beberapa keuntungan, antara lain:

- Mengurangi pengaruh kata-kata umum yang kurang memiliki nilai informasi.
- Menonjolkan kata-kata yang memiliki makna penting dalam menentukan sentimen.
- Menghasilkan representasi numerik yang lebih informatif dibandingkan metode frekuensi kata biasa (*Bag of Words*).

Dalam penelitian ini, TF-IDF membantu model mengidentifikasi kata-kata yang sering muncul pada ulasan positif maupun negatif sehingga proses klasifikasi sentimen dapat dilakukan dengan lebih baik.

3. Penggunaan parameter `ngram_range=(1,2)`

ada proses ekstraksi fitur menggunakan TF-IDF, diterapkan parameter **`ngram_range=(1,2)`**. Parameter ini memungkinkan sistem untuk menggunakan kombinasi **unigram** (satu kata) dan **bigram** (dua kata berurutan) sebagai fitur dalam proses klasifikasi.

Penggunaan unigram memungkinkan model mengenali kata-kata tunggal seperti:

- Bagus
- Cepat
- Mahal
- Buruk

Sedangkan penggunaan bigram memungkinkan model mengenali kombinasi kata seperti:

- Sangat bagus
- Tidak puas
- Pengiriman cepat
- Kualitas buruk

Penerapan bigram memberikan keuntungan karena mampu menangkap konteks yang tidak dapat dipahami jika hanya menggunakan unigram. Sebagai contoh,

kata "bagus" memiliki sentimen positif, namun pada kalimat "tidak bagus" maknanya berubah menjadi negatif. Dengan adanya fitur bigram, model dapat mengenali frasa "tidak bagus" sebagai indikator sentimen negatif.

Oleh karena itu, penggunaan parameter $\text{ngram_range}=(1,2)$ meningkatkan kemampuan model dalam memahami hubungan antar kata sehingga menghasilkan klasifikasi yang lebih akurat dibandingkan hanya menggunakan unigram saja.

4. Algoritma Naive Bayes

Algoritma **Naive Bayes** digunakan sebagai metode klasifikasi sentimen karena memiliki performa yang baik pada data teks dan proses komputasi yang relatif sederhana. Algoritma ini bekerja berdasarkan konsep probabilitas dengan menghitung kemungkinan suatu dokumen termasuk ke dalam kelas tertentu berdasarkan kata-kata yang terdapat di dalamnya.

Pada penelitian ini, Naive Bayes digunakan untuk mengklasifikasikan ulasan menjadi sentimen positif dan negatif. Setiap kata yang muncul dalam ulasan akan dihitung probabilitas kemunculannya pada masing-masing kelas sentimen. Kelas dengan nilai probabilitas tertinggi akan menjadi hasil prediksi model.

Beberapa alasan penggunaan Naive Bayes dalam penelitian ini adalah:

1. Memiliki proses pelatihan yang cepat.
2. Efektif untuk klasifikasi data teks berdimensi tinggi.
3. Mampu bekerja dengan baik meskipun jumlah data pelatihan tidak terlalu besar.
4. Mudah diimplementasikan dan memiliki kebutuhan komputasi yang rendah.

Hasil penerapan Naive Bayes menunjukkan bahwa algoritma mampu memanfaatkan fitur-fitur yang dihasilkan oleh TF-IDF untuk membedakan ulasan positif dan negatif secara efektif. Kombinasi antara preprocessing yang baik, ekstraksi fitur menggunakan TF-IDF dengan $\text{ngram_range}=(1,2)$, serta algoritma Naive Bayes menghasilkan model analisis sentimen yang mampu memberikan performa klasifikasi yang optimal.

SIMPULAN

Berdasarkan hasil penelitian, algoritma **Multinomial Naive Bayes** mampu mengklasifikasikan sentimen ulasan produk Tokopedia ke dalam kategori positif dan negatif dengan baik. Tahap **preprocessing** dan ekstraksi fitur menggunakan **TF-IDF** berhasil meningkatkan kualitas data sehingga mendukung proses klasifikasi. Hasil pengujian menunjukkan model memperoleh **akurasi sebesar 83%**, yang menunjukkan bahwa metode yang digunakan cukup efektif dalam melakukan analisis sentimen pada ulasan pelanggan Tokopedia. Penelitian ini diharapkan dapat menjadi referensi dalam membantu penjual maupun platform e-commerce memahami opini pelanggan untuk meningkatkan kualitas produk dan layanan.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Tuhan Yang Maha Esa, dosen pembimbing, Program Studi Ilmu Komputer Fakultas Teknik dan Informatika Universitas Widyagama Mahakam Samarinda, serta seluruh pihak yang telah memberikan dukungan, bimbingan, dan bantuan selama proses penelitian dan penyusunan artikel ini. Penulis berharap penelitian ini dapat memberikan manfaat bagi pengembangan ilmu pengetahuan, khususnya pada bidang analisis sentimen dan *machine learning*.

DAFTAR PUSTAKA

- [1] K. Kunci, "Analisis Sentimen E-Commerce di Indonesia dengan Algoritma Naive Bayes (Studi Kasus Pada Platform Shopee)," vol. 01, no. 04, pp. 109–114, 2026.
- [2] R. Maulana, P. Hertaryawan, M. Raihan, and I. Santoso, "KOMPARASI ALGORITMA NAIVE BAYES DAN K- NEAREST NEIGHBOR PADA ANALISIS SENTIMEN," vol. 17, no. 2, pp. 177–189, 2023.
- [3] A. Gondowastadmojo, R. Kusumastuti, and S. Indonesia, "Analisa sentimen ulasan di tokopedia dengan metode naive bayes," no. November, pp. 369–382, 2024.
- [4] B. Z. Ramadhan, I. Riza, and I. Maulana, "Analisis Sentimen Ulasan Pada Aplikasi E-Commerce Dengan Menggunakan Algoritma Naive Bayes," vol. 6, no. 2, pp. 220–225, 2022.
- [5] N. Tahun, N. A. Sivi, I. Mualim, and M. T. Kussofyan, "Penerapan Algoritma Naive Bayes untuk Analisis Sentimen Ulasan Produk model analisis sentimen pada ulasan produk e-commerce berbahasa Indonesia dan mengukur " kotor " (noisy), diperlukan serangkaian tahapan text preprocessing . Tahapan ini krusial dan," 2023.
- [6] F. A. Lubis et al., "IMPLEMENTASI BIG DATA DALAM ANALISIS SENTIMEN ULASAN PENGGUNA TOKOPEDIA BERBASIS APLIKASI WEB MENGGUNAKAN METODE NAIVE BAYES," pp. 314–321, 2026.
- [7] R. Ferdiansyah et al., "Analisis Sentimen Dalam Tokopedia Terhadap Ulasan Pelanggan Menggunakan Metode Naive Bayes," vol. 9, pp. 24–29.
- [8] R. T. Nagur, D. Alamsyah, M. G. Tauran, Y. Nuryamin, and A. Priyatna, "Perbandingan Kinerja Naive Bayes dan Decision Tree untuk Analisis Sentimen Ulasan Produk Online di Shopee dan Tokopedia," vol. 2024, pp. 124–132, 2025.
- [9] H. Barus, I. N. Fajri, and Y. Pristyanto, "Sentiment Classification Analysis of Tokopedia Reviews Using TF-IDF , SMOTE , and Traditional Machine Learning Models," vol. 9, no. 5, pp. 2552–2561, 2025.
- [10] C. C. Kirana and N. R. Oktadini, "Analisis Sentimen Naive Bayes dengan TF-IDF dan 10-Fold pada Ulasan Aplikasi X," vol. 7, pp. 523–535, 2025, doi: 10.30865/json.v7i2.9007.
- [11] J. Edukasi, B. Gunawan, H. S. Pratiwi, and E. E. Pratama, "Sistem Analisis Sentimen pada Ulasan Produk Menggunakan Metode Naive Bayes," vol. 4, no. 2, pp. 113–118, 2018.
- [12] A. A. Firdauzi and S. T. Annaztasya, "Systematic Literature Review : Analisis Dan Prediksi Perilaku Konsumen E-Commerce Menggunakan Machine Learning," vol. 2, no. 3, pp. 455–465, 2026.
- [13] B. Siswoyo, N. Azka, and P. Utomo, "Pemanfaatan Machine learning untuk Klasifikasi Sentimen Pelanggan pada Media Sosial," vol. 1, no. 1, pp. 29–34, 2025.
- [14] T. M. A. Admira and M. Saefudin, "MODEL SENTIMENT ANALYSIS BERBASIS MACHINE LEARNING UNTUK DATA GENZ-CAREER ASPIRATION MENGGUNAKAN FLASK DAN NAIVE BAYES (Machine Learning Based Sentiment Analysis Model For Genz-Career Aspiration Data Using Flask and Naive Bayes)," vol. 9, no. 1, pp. 25–39, 2025, doi: 10.52362/jisicom.v9i1.1880.
- [15] A. M. Rangkuti, H. Mustafa, A. Guna, and M. Tarihoran, "Klasifikasi Sentimen Ulasan Produk E-Commerce Menggunakan Naive Bayes," vol. 1, pp. 48–58, 2026.
- [16] E. K. Fitri, A. H. Anshor, and M. S. Anwar, "Perbandingan Metode Machine Learning Klasik (Naive Bayes dan SVM) dan IndoBERT untuk Analisis Sentimen Ulasan Produk pada

- Platform Tokopedia,” vol. 11, no. 6, pp. 4806–4816, 2026.
- [17] K. S. U. E-commerce and M. N. Bayes, “Klasifikasi sentimen ulasan e-commerce menggunakan naïve bayes , lightgbm, dan gru”.
- [18] K. Dan and P. Ulasan, “Klasifikasi dan prediksi ulasan e-commerce menggunakan algoritma naïve bayes 1-4),” vol. 9, no. 1, pp. 207–217, 2025.
- [19] A. S. Wanda, “Implementasi Naïve Bayes Untuk Klasifikasi Sentimen Review Tokopedia Berbasis Web Dashboard Implementation of Naïve Bayes for Tokopedia Review Sentiment Classification Based on Web Dashboard Jurnal Armada Informatika,” vol. 10, pp. 103–112, 2026.
- [20] S. Ardiawal, E. Rahmi, and E. Kurniawaty, “Analisis Sentimen Ulasan untuk Menentukan Kualitas Produk Terbaik E- Commerce (Tokopedia) Menggunakan Algoritma Naive Bayes,” no. November, 2025.