

Analisis Sentimen Produk Kosmetik Menggunakan Naive Bayes dan SVM

Cosmetic Product Sentiment Analysis Using Naive Bayes and SVM

Elisabeth Jesica Mirino¹, Rambu Aya², Agnes Niaviyane³, Olivia Raku⁴

¹²³⁴Program Studi Teknik Informatika, Fakultas Teknik, Universitas Papua

Jl. Gunung Salju, Amban, Manokwari, Papua Barat

Telp : (0986) +6281241026393

Email: ¹202365029@student.unipa.ac.id, ²202365030@student.unipa.ac.id, ³202365039@student.unipa.ac.id,
⁴202365054@student.unipa.ac.id

Abstract

The growth of e-commerce in Indonesia has increased online sales of cosmetics. Consumer reviews on e-commerce platforms are an important source of information for purchasing decisions. However, the large volume of reviews makes manual analysis inefficient. This study aims to analyze consumer review sentiment towards cosmetics on e-commerce platforms using the Naive Bayes and Support Vector Machine (SVM) algorithms and to compare their performance. The sentiment dataset was obtained from Kaggle.com, with 200 samples per class (600 data points in total). After text preprocessing, 599 valid data points remained for classification with the Naive Bayes and SVM algorithms. Text preprocessing included case folding, cleaning, tokenization, and stopword removal. Feature extraction was performed using TF-IDF, with data divided into 80% training and 20% testing. The results showed that the Naive Bayes algorithm achieved 62.50% accuracy with an F1-score of 62.80%, while the SVM (RBF kernel) achieved 60.00% accuracy with an F1-score of 60.00%. Naive Bayes outperformed the SVM across all evaluation metrics in the sentiment classification of Indonesian-language cosmetic product reviews.

Keywords: Sentiment analysis; Cosmetics; e-commerce; Naive Bayes; Support Vector Machines; TF-IDF.

Abstrak

Perkembangan e-commerce di Indonesia mendorong meningkatnya transaksi penjualan produk kosmetik secara daring. Ulasan konsumen yang ditinggalkan pada platform e-commerce menjadi sumber informasi penting dalam pengambilan keputusan pembelian. Namun, volume ulasan yang sangat besar membuat analisis secara manual menjadi tidak efisien. Penelitian ini bertujuan untuk menganalisis sentimen ulasan konsumen terhadap produk kosmetik di e-commerce menggunakan algoritma Naive Bayes dan Support Vector Machine (SVM), serta membandingkan performa kedua algoritma. Dataset sentimen diperoleh dari kaggle.com dengan total masing-masing kelas berjumlah 200 sampel (total 600 data), yang setelah di-preprocessing teks tersisa 599 data valid untuk proses klasifikasi menggunakan algoritma Naive Bayes dan SVM. Preprocessing teks meliputi case folding, cleaning, tokenisasi, dan penghapusan stopword. Ekstraksi fitur dilakukan menggunakan TF-IDF dengan pembagian data 80% training dan 20% testing. Hasil penelitian menunjukkan bahwa algoritma Naive Bayes memperoleh akurasi 62,50% dengan F1-Score 62,80%, sedangkan SVM (kernel RBF) memperoleh akurasi 60,00% dengan F1-Score 60,00%. Naive Bayes mengungguli SVM pada seluruh matriks evaluasi dalam klasifikasi sentimen ulasan produk kosmetik berbahasa Indonesia.

Kata Kunci: sentiment analysis; cosmetics; e-commerce; naive bayes; support vector machine; TF-IDF.

1. Pendahuluan

Perkembangan teknologi informasi telah mendorong peningkatan pesat pada platform e-commerce di Indonesia. Dengan adanya e-commerce yang mendukung transaksi antara penjual dan pembeli secara daring tanpa dibatasi oleh ruang dan waktu, hal ini memberikan kemudahan pada kedua belah pihak [1]. Salah satu

bagian yang mengalami kemajuan penting dalam e-commerce adalah produk kosmetik, di mana konsumen tidak hanya membeli secara daring, tetapi juga aktif dalam memberikan ulasan untuk membagikan pengalaman pengguna. Ulasan tersebut dijadikan oleh konsumen lain sebagai informasi penting untuk mengambil keputusan, di mana ulasan tersebut berisi penilaian terhadap kualitas, manfaat, serta kekurangan suatu produk [2].

Namun, dengan bertambahnya jumlah pengguna e-commerce, volume ulasan yang didapatkan juga semakin banyak dan tidak terstruktur. Keadaan ini menyulitkan proses analisis bilamana dilakukan secara manual, sehingga dibutuhkan pendekatan yang bisa mengolah data teks secara otomatis. Dalam hal ini, analisis sentimen digunakan untuk mengidentifikasi dan mengelompokkan pendapat konsumen ke dalam beberapa kategori tertentu, seperti positif, negatif, dan netral [10].

Beberapa metode yang sering digunakan dalam analisis sentimen adalah Naive Bayes (NB) dan Support Vector Machine (SVM). Naive Bayes dikenal sebagai metode yang sederhana, cepat, dan efektif dalam mengolah data teks, serta mampu menghasilkan performa yang baik dalam berbagai penelitian [4]. Sementara itu, Support Vector Machine memiliki kemampuan dalam menangani data berdimensi tinggi dan sering menghasilkan tingkat kinerja yang lebih tinggi dalam klasifikasi teks [3].

Pada penelitian sebelumnya, sudah banyak analisis sentimen menggunakan algoritma Naive Bayes dan Support Vector Machine. Salah satunya yang dilakukan oleh D. A. WP, J. D. Firizqi, dan Z. A. Amalia yang berjudul “Analisis Sentimen Produk Skincare Something Niacinamide di Female Daily dengan Naive Bayes Classifier” yang menunjukkan bahwa metode Naive Bayes mampu mengklasifikasikan sentimen dengan baik pada data ulasan kosmetik [1], sedangkan penelitian yang menggunakan Support Vector Machine pada ulasan produk skincare di e-commerce bisa mencapai akurasi sebesar 94% [3]. Sedangkan penelitian yang dilakukan oleh J. A. Putra, A. Dharmawan, dan J. Gondohanindjo dengan judul “Sentimen Analisis Aplikasi Digitalent Mobile Menggunakan Naive Bayes dan SVM dengan Ekstraksi Fitur TF-IDF” menunjukkan bahwa metode Naive Bayes bisa mencapai akurasi rata-rata sampai 95% pada data e-commerce [11].

Penelitian perbandingan yang dilakukan oleh J. A. Putra, A. Dharmawan, dan J. Gondohanindjo yang berjudul “Sentimen Analisis Aplikasi Digitalent Mobile Menggunakan Naive Bayes dan SVM dengan Ekstraksi Fitur TF-IDF” juga menampilkan bahwa Support Vector Machine sering mempunyai kinerja lebih tinggi dibandingkan Naive Bayes, dengan akurasi sebanyak 98.7% dibandingkan 88.66% [5]. Namun, pada penelitian lain ditemukan bahwa Naive Bayes justru bisa memberikan hasil yang lebih baik dibandingkan dengan SVM, tergantung pada jenis data yang digunakan [9]. Hal tersebut menunjukkan bahwa belum ada metode yang secara tepat dan sempurna dalam analisis sentimen, karena performa algoritma dipengaruhi oleh karakteristik

data, teknik preprocessing, serta metode ekstraksi fitur yang digunakan [13].

Selain itu, sebagian penelitian menggunakan dataset umum seperti media sosial dan aplikasi, sehingga belum secara jelas mempresentasikan ciri ulasan produk kosmetik pada e-commerce karena sebagian data berasal dari platform Twitter [2]. Sementara itu, terdapat beberapa peneliti yang hanya menggunakan salah satu metode tanpa melakukan perbandingan langsung dalam satu penelitian [1]. Oleh karena itu, penelitian ini memiliki kebaruan dengan melakukan perbandingan langsung antara algoritma Naive Bayes dan Support Vector Machine pada database yang sama. Selain itu, penelitian ini juga menggunakan data ulasan produk kosmetik berbahasa Indonesia dari platform e-commerce dengan tahapan preprocessing berupa tokenisasi, stemming, dan pembobotan TF-IDF. Sehingga memberikan hasil yang lebih nyata dan relevan. Kontribusi penelitian ini adalah memberikan analisis perbandingan cara kerja dua algoritma klasifikasi, yaitu Naive Bayes dan Support Vector Machine, pada data ulasan produk kosmetik di e-commerce.

Tujuan dari penelitian ini adalah menganalisis sentimen ulasan konsumen terhadap produk kosmetik di e-commerce menggunakan algoritma Naive Bayes dan Support Vector Machine dan secara khusus membandingkan kinerja algoritma Naive Bayes dan Support Vector Machine pada analisis sentimen ulasan produk kosmetik di platform e-commerce dalam satu konteks penelitian yang sama. Hasil penelitian ini diharapkan dapat memberikan gambaran mengenai metode yang lebih ideal dalam klasifikasi sentimen, khususnya pada data ulasan produk kosmetik di e-commerce. Penelitian ini juga memiliki keterbatasan dalam jumlah data yang digunakan, yaitu sebanyak 599 ulasan yang diperoleh setelah preprocessing dari database awal, yaitu 1000 data, dan hanya menggunakan dua algoritma klasifikasi. Karena itu, hasil yang didapat bisa dikembangkan lebih lanjut oleh penelitian selanjutnya.

2. Tinjauan Pustaka

2.1. Analisis Sentimen

Analisis sentimen adalah cara menganalisis ulasan pengguna tentang bagaimana sebuah sentimen diekspresikan menggunakan teks dan bagaimana sentimen tersebut bisa dikategorikan sebagai sentimen positif, sentimen negatif, atau sentimen netral [20].

2.2. Naive Bayes

Naive Bayes menggunakan prinsip Teorema Bayes untuk menghitung peluang kemungkinan suatu data tergolong dalam kelas tertentu berdasarkan data yang sudah ada sebelumnya. Data bisa berupa ulasan sentimen dengan menentukan sebuah ulasan sebagai sentimen negatif, sentimen positif, atau sentimen netral [21]. Bentuk umum teorema bayes adalah sebagai berikut:

$$P(H|X) = \frac{P(X|H)P(H)}{P(X)} \quad (1)$$

Keterangan:

X : data yang belum memiliki label kelas

H : hipotesis X termasuk kelas tertentu

P(H|X): peluang hipotesis H (prior Probability)

P(H) : peluang X berdasar kondisi hipotesis H

P(X) : peluang dari X

Proses penyelesaian Naive Bayes dilakukan dengan menghitung peluang P(H), yaitu peluang X berdasarkan kondisi hipotesis H, H merupakan kelas sentimen positif, negatif atau netral, yang hasilnya dikalikan antar kata untuk mendapatkan nilai P(H|X) dari suatu kelas. Kemudian, P(X) yaitu peluang kemunculan X secara keseluruhan, yang memiliki nilai yang sama untuk semua kelas sehingga tidak ikut diperhitungkan karena tidak mengubah hasil akhir. Terakhir, P(H|X) yaitu peluang hipotesis H yang menghitung setiap kelas dari data X yang belum diketahui kelasnya, kemudian akan dimasukkan ke dalam kelas sentimen dengan nilai P(H|X) tertinggi. Maka, data tersebut digolongkan sebagai hasil akhir prediksi.

2.3. Support Vector Machine (SVM)

Support Vector Machine (SVM) adalah suatu metode klasifikasi yang digunakan untuk memisahkan data ke dalam kelas tertentu, dengan cara mencari batas pemisah agar setiap kelompok data dapat dibedakan dengan jelas. Persamaan SVM dapat dilihat pada [22].

$$\sum_{i=1}^{ns} a_i y_i \rightarrow x_1 \cdot x_d + b \quad (2)$$

Dimana:

ns = Jumlah support vector
ai = Nilai bobot yang ditentukan pada setiap data
yi = Kelas data
i = Variable support vector
d = Data yang akan diklasifikasi
b = Nilai error atau bias

Proses penyelesaian SVM dilakukan dengan menjumlahkan hasil perkalian antara ai (nilai bobot yang ditentukan pada setiap data), yi (kelas data), dan xi (variabel support vector) terhadap data yang mau diklasifikasikan, kemudian dijumlahkan semua sebanyak jumlah support vector yang ada. Lalu hasilnya ditambah dengan nilai bias untuk mengoreksi hasil perhitungan. Hasil akhir menentukan posisi data masuk ke sentimen mana.

2.4. TF- IDF

Term Frequency - Inverse Document Frequency atau TF-IDF merupakan metode yang umum digunakan untuk menghitung bobot setiap kata dalam sebuah dokumen, dengan melihat seberapa sering kata tersebut muncul dan seberapa penting kata tersebut dalam dokumen [9]. Adapun rumus pembobotan data TF-IDF sebagai berikut:

$$W(t,d) = TF(t,d) \times \log\left(\frac{N}{DF(t)}\right) \quad (3)$$

Keterangan:

Wt,d: Bobot TF- IDF

t f t,d : Jumlah Frekuensi kata

i d f t : Jumlah inverse yang menggambarkan seberapa jarang suatu kata muncul pada dokumen

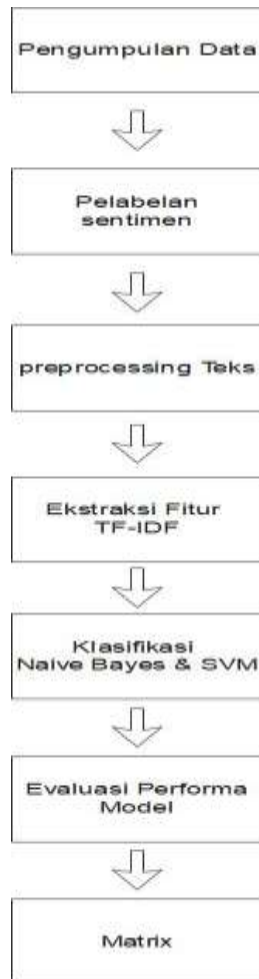
DF(t) : Jumlah frekuensi dokumen tiap kata

N : Jumlah total dokumen

Proses penyelesaian TF-IDF dilakukan dengan menghitung TF(t,d), yaitu seberapa sering suatu kata muncul dalam sebuah dokumen. Kemudian IDF dihitung dengan membagi jumlah total dokumen dengan jumlah dokumen yang mengandung kata tersebut, DF(t), lalu hasilnya akan di log kan agar nilai tidak terlalu besar. Kemudian kedua nilai tersebut dikalikan dan hasilnya disebut W(t,d) atau bobot TF-IDF.

3. Metodologi Penelitian

Metode penelitian ini dilakukan melalui beberapa tahapan sistematis sebagaimana ditunjukkan pada Gambar 1. Setiap tahapan dirancang untuk memastikan proses analisis sentimen berjalan secara terstruktur dan menghasilkan model klasifikasi yang akurat.



Gambar 1 metode penelitian

Tahapan pertama adalah pengumpulan **dataset**, yaitu mengumpulkan data teks berupa komentar atau ulasan dari objek penelitian yang relevan, seperti media sosial atau platform ulasan produk. Dataset sentimen diperoleh dari [kaggle.com](https://www.kaggle.com) dengan total awal 1000 data yang mengalami ketidakseimbangan kelas, di mana label negatif hanya 200 sampel sementara label positif dan netral mendominasi. Untuk menyeimbangkannya, dilakukan random undersampling pada kelas mayoritas sehingga masing-masing kelas berjumlah 200 sampel (total 600 data). Setelah preprocessing teks, tersisa 599 data valid untuk proses klasifikasi menggunakan

algoritma Naive Bayes dan SVM. Data dikumpulkan menggunakan teknik scraping atau unduhan manual sesuai kebutuhan penelitian. Selanjutnya dilakukan **preprocessing teks**, yaitu proses pembersihan dan normalisasi data teks menggunakan teknik tokenisasi, penghapusan stopword, dan stemming dengan bantuan library Natural Language Processing seperti NLTK atau Sastrawi khusus untuk bahasa Indonesia, guna memastikan data yang diproses bebas dari noise dan tidak relevan.

Tahapan berikutnya adalah **pelabelan sentimen**, di mana setiap data teks diberi label kategori positif, negatif, atau netral menggunakan pendekatan lexicon-based atau anotasi manual oleh peneliti. Setelah data berlabel, dilakukan **ekstraksi fitur TF-IDF** (Term Frequency-Inverse Document Frequency) untuk mengubah data teks menjadi representasi vektor numerik berdasarkan bobot kepentingan setiap kata dalam dokumen menggunakan library scikit-learn. Tahapan terakhir adalah **klasifikasi dan evaluasi**, di mana model machine learning seperti Naive Bayes, SVM, atau algoritma lainnya dilatih menggunakan data hasil ekstraksi fitur, kemudian dievaluasi menggunakan metrik akurasi, presisi, recall, dan F1-score untuk mengukur performa model yang dibangun.

3.1. Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan data sekunder berupa ulasan konsumen yang diperoleh dari platform e-commerce seperti *Femaledaily.com*. Data ulasan tersebut mencerminkan pengalaman pengguna terhadap produk yang dibeli, baik dari segi kualitas produk, kecepatan, pengiriman, maupun pelayanan penjual

Data yang digunakan dalam penelitian ini terdiri dari teks ulasan yang telah dikategorikan ke dalam tiga label sentimen, yaitu positif, netral, dan negatif. proses pengumpulan data dilakukan dengan metode pengambilan data secara tidak langsung (data sekunder), baik melalui scraping maupun data sat publik yang relevan dengan topik penelitian[4]

Jumlah data yang digunakan sebanyak ± 1000 data ulasan. Distribusi data dibuat seimbang untuk menghindari bias dalam proses pelatihan model, dengan rincian sekitar 200 data berlabel positif, 200 data berlabel netral, dan 199 data berlabel negatif. Keseimbangan data ini penting untuk memastikan bahwa model klasifikasi tidak cenderung pada salah satu kelas tertentu.

Data yang digunakan sebagai objek penelitian untuk dianalisis lebih lanjut. Dalam

penelitian analisis sentimen, dataset biasanya berupa teks ulasan pengguna yang mengandung opini terhadap suatu produk. Dataset teks menjadi komponen utama dalam proses *text mining* karena berfungsi sebagai sumber informasi yang akan diolah untuk menemukan pola sentimen. Oleh karena itu, dataset yang digunakan harus relevan dengan tujuan penelitian agar hasil klasifikasi dapat merepresentasikan kondisi sebenarnya.[4]

3.2. Pelabelan Sentimen

Pelabelan sentimen merupakan proses pemberian label terhadap setiap ulasan ke dalam kategori tertentu, yaitu positif, netral, dan negatif, berdasarkan isi, makna, serta konteks kalimat yang terkandung di dalamnya. Proses ini umumnya dilakukan secara manual oleh peneliti untuk memastikan tingkat keakuratan label yang tinggi, terutama pada data berbahasa alami yang sering mengandung ambiguitas, ironi, atau makna ganda. Dalam pelaksanaannya, peneliti perlu memahami konteks kalimat secara menyeluruh agar hasil pelabelan sesuai dengan maksud sebenarnya dari penulis ulasan.

Tujuan utama dari pelabelan sentimen adalah untuk menyediakan data latih (*training data*) yang akan digunakan dalam proses pembelajaran model klasifikasi. Data yang telah diberi label ini akan menjadi acuan bagi algoritma dalam mengenali pola dan karakteristik dari masing-masing kategori sentimen. Oleh karena itu, kualitas pelabelan sangat berpengaruh terhadap performa model yang dihasilkan. Pelabelan yang tidak konsisten atau kurang tepat dapat menyebabkan model mengalami kesalahan dalam proses klasifikasi. Dengan demikian, pelabelan sentimen menjadi tahap yang sangat penting karena menentukan arah pembelajaran algoritma dalam membedakan kategori sentimen secara akurat.[4]

3.3. Preprocessing Teks

Preprocessing teks merupakan tahapan awal yang dilakukan untuk membersihkan dan menyiapkan data sebelum proses analisis. Tahapan ini meliputi case folding (mengubah seluruh huruf menjadi kecil), cleaning (menghapus tanda baca, angka, dan karakter yang tidak diperlukan), tokenizing (memecah kalimat menjadi kata-kata), stopword removal (menghapus kata umum yang tidak memiliki makna penting), serta stemming (mengubah kata ke bentuk dasar). Proses ini bertujuan untuk menghilangkan noise, melakukan normalisasi teks, serta menyederhanakan struktur kata sehingga kualitas data meningkat dan

memudahkan algoritma dalam mengenali pola, yang pada akhirnya dapat meningkatkan performa model klasifikasi.[4]

3.4. Ekstraksi Fitur TF-IDF

Ekstraksi fitur merupakan tahapan penting dalam pengolahan data teks yang bertujuan untuk mengubah data dalam bentuk teks menjadi representasi numerik agar dapat diproses oleh algoritma *machine learning*. Hal ini dilakukan karena komputer tidak dapat secara langsung memahami teks, sehingga diperlukan transformasi ke dalam bentuk angka. Dalam penelitian ini digunakan metode TF-IDF (*Term Frequency–Inverse Document Frequency*), yaitu teknik pembobotan kata yang mengukur tingkat kepentingan suatu kata dalam dokumen

TF-IDF bekerja dengan mengkombinasikan *Term Frequency* (TF), yaitu frekuensi kemunculan kata dalam suatu dokumen, dan *Inverse Document Frequency* (IDF), yaitu tingkat keunikan kata dalam keseluruhan dokumen. Kata yang sering muncul dalam satu dokumen tetapi jarang muncul pada dokumen lain akan memiliki bobot lebih tinggi, sedangkan kata yang umum muncul di banyak dokumen akan memiliki bobot lebih rendah. Dengan metode ini, setiap dokumen direpresentasikan dalam bentuk vektor numerik yang mampu menonjolkan kata-kata penting, sehingga dapat membantu meningkatkan kinerja model klasifikasi dalam mengidentifikasi sentimen secara lebih akurat.[4]

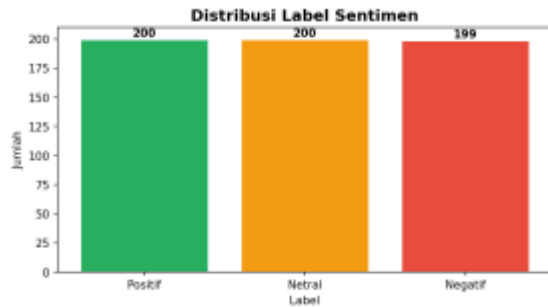
3.5. Klasifikasi dan Evaluasi

Proses klasifikasi dilakukan menggunakan algoritma Naive Bayes dan Support Vector Machine (SVM) dengan kernel RBF. Proses pembagian data menggunakan rasio 80:20, di mana porsi lebih besar digunakan untuk melatih model dan sisanya untuk validasi hasil klasifikasi. Evaluasi model dilakukan menggunakan confusion matrix serta metrik akurasi, presisi, recall, dan F1-score untuk mengukur dan membandingkan performa kedua metode dalam klasifikasi sentimen.

Salah satu metode yang banyak digunakan adalah Naive Bayes karena kesederhanaannya dan kemampuannya dalam mengolah data teks. Algoritma Naive Bayes mampu memberikan hasil klasifikasi yang cukup baik dalam analisis sentimen.[4]

4. HASIL DAN PEMBAHASAN

4.1. Hasil Preprocessing



Gambar 2. Distribusi label

Setelah melalui tahap preprocessing, sebanyak 599 ulasan berhasil diproses menjadi data teks yang bersih dan siap diekstraksi fiturnya. Distribusi data training terdiri dari 160 ulasan positif, 160 ulasan netral, dan 159 ulasan negatif dengan total ada 479 ulasan, sedangkan data testing terdiri dari 40 ulasan per kelas dengan total ada 120 ulasan. Proses tokenisasi menghasilkan rata-rata 12,5 token per ulasan sebelum stopwords removal dan 8,3 token setelah stopwords removal. Dari hasil ini juga diketahui setiap kata yang sering muncul pada sentimen positif antara lain “bagus”, “cocok”, “lembab”, dan “cerah”, sedangkan kata yang sering muncul pada sentimen negatif antara lain “iritasi”, “lengket”, “kecewa”, dan “tidak cocok”.

3.2. Hasil Klasifikasi dan Evaluasi

Tabel 1. Perbandingan Performa Naive Bayes dan SVM

Algrt	Akurasi	Presisi	Recall	F1-Score
NB	63%	64%	63%	63%
SVM	60%	61%	60%	60%

Berdasarkan Tabel 1, algoritma Naive Bayes menunjukkan performa lebih baik dibandingkan dengan SVM pada seluruh matriks evaluasi. Naive Bayes memperoleh akurasi 62,50% dan F1-Score 62,80%, sedangkan SVM(RBF) memperoleh akurasi 60,00% dan F1-Score 60,00%. Dari hasil tersebut didapatkan selisih akurasi 2,50%. Hal ini menunjukkan Naive Bayes lebih sesuai untuk klasifikasi sentimen ulasan kosmetik berbahasa Indonesia pada dataset ini.

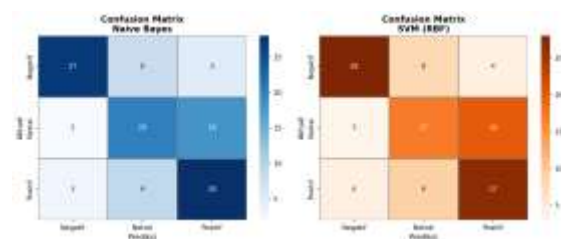
Jurnal Ilmiah Binary STMIK Bina Nusantara Jaya

Vol.8 No. 01 Tahun 2026, ISSN : 2657– 2117 | DOI 10.52303/jb.v8i1.217

Tabel 2. Performa Per Kelas Sentimen

Clas	NB -P	NB -R	NB- F1	SVM- P	SVM- R	SVM- F1
Negatif	84 %	68 %	75%	80%	70%	75%
Netral	54 %	50 %	52%	50%	43%	46%
Positif	55 %	70 %	62%	53%	68%	59%

Berdasarkan Tabel 2, kedua algoritma menunjukkan performa terbaik pada klasifikasi kelas negatif dengan F1-Score Naive Bayes sebesar 75,0% dan SVM sebesar 74,7%. Hasil ini diperoleh dari karakteristik ulasan negatif yang mengandung kata-kata emosional yang mudah dikenali oleh model, seperti “iritasi”, “lengket”, “kecewa”, dan “tidak cocok”. Kata-kata tersebut menunjukkan pola distribusi yang sangat berbeda antara kelas negatif dan kelas lainnya, sehingga mempermudah model dalam mendeteksi ulasan negatif. Sebaliknya, kelas netral menunjukkan nilai terendah pada kedua algoritma, yaitu 51,9% untuk Naive Bayes dan 45,9% untuk SVM. Hal ini disebabkan oleh ulasan netral yang cenderung tidak mengandung ekspresi emosional yang kuat sehingga lebih sulit dibedakan oleh model antara ulasan positif dan negatif.



Gambar 3. Confusion Matrix

Sebagaimana ditampilkan pada Gambar 3, confusion matrix menunjukkan bahwa kesalahan klasifikasi terbesar terjadi pada kelas Netral yang sering salah diprediksi sebagai positif, yaitu 18 sampel pada Naive Bayes dan 20 sampel pada SVM. Keunggulan Naive Bayes dibanding SVM pada penelitian ini disebabkan oleh kemampuannya menghitung probabilitas kemunculan kata kunci sentimen secara efektif pada dataset berukuran sedang [16]. Dengan ukuran data training yang terbatas, 479 sampel, Naive Bayes cenderung lebih

stabil dibandingkan dengan SVM yang memerlukan optimasi hyperplane. Hal ini konsisten dengan penelitian [17] yang menemukan bahwa Naive Bayes lebih unggul dibandingkan dengan SVM pada dataset berukuran kecil hingga menengah pada ulasan e-commerce berbahasa Indonesia.



Gambar 4. Perbandingan Performa Naive Bayes dan SVM

Gambar 4 menunjukkan bahwa pada kedua model yang bernilai antara 60%–63% masih terdapat peluang untuk peningkatan. Faktor-faktor yang berkontribusi pada keterbatasan akurasi ini antara lain: (1) tidak ditetapkannya stemming bahasa Indonesia yang dapat mengurangi variasi morfologis kata, (2) ukuran dataset yang terbatas, (3) ambiguitas semantik pada teks ulasan informal bahasa Indonesia yang menggunakan banyak slang dan singkatan. Penggunaan metode deep learning berbasis konteks seperti IndoBet berpotensi memberikan peningkatan akurasi yang signifikan pada penelitian lanjutan [18, 19].

5. Kesimpulan

5.1. Simpulan

Penelitian ini menunjukkan bahwa Naive Bayes lebih sesuai dibandingkan dengan SVM dengan RBF dalam klasifikasi ulasan sentimen kosmetik pada dataset ini. Dengan pembagian data training sebesar 80% dan data testing sebesar 20%, Naive Bayes menghasilkan akurasi sebesar 62,50% dan F1-Score sebesar 62,83%, sedangkan SVM mencapai tingkat akurasi sebesar 60,00% dan F1-Score sebesar 59,98%. Kemudian performa per kelas, kelas negatif mengandung kata-kata emosional yang mudah dikenali oleh model, seperti “iritasi”, “lengket”, “kecewa”, dan “tidak cocok”, dan kelas netral cenderung tidak mengandung ekspresi emosional yang kuat sehingga lebih sulit dibedakan oleh model antara ulasan positif dan negatif. Adapun penelitian terdahulu yang melakukan perbandingan Naive

Bayes dan Support Vector Machine diantaranya penelitian yang dilakukan oleh (T.P.R.Sanjaya & A.Fauzi,2023) untuk analisis sentimen ulasan pada e-commerce shopee menghasilkan akurasi Naive bayes sebesar 85% dan SVM sebesar 81%. Selanjutnya, penelitian yang dilakukan oleh (J.A.Putra & Dharmawan,2024) untuk analisis sentimen aplikasi digital ent mobile menghasilkan akurasi Naive Bayes sebesar 85,7% dan SVM sebesar 82,14%. Adapun (R.Vincent & I.Maulana,2023), analisis sentimen multiclass di Twitter menghasilkan akurasi Naive Bayes sebesar 72,02%, sedangkan SVM menghasilkan akurasi sebesar 62,38%. Hasil pengujian menunjukkan Naive Bayes memberikan kinerja lebih baik dibandingkan dengan SVM dikarenakan ukuran dataset sedang, yaitu 479 sampel, dan kemampuannya menghitung probabilitas kemunculan kata kunci sentimen secara efektif. Naive Bayes cenderung lebih stabil dibandingkan dengan SVM yang memerlukan optimasi hyperplane. Hal ini menunjukkan bahwa Naive Bayes lebih unggul pada dataset yang kecil hingga menengah.

5.2. Saran

Penelitian selanjutnya disarankan untuk menggunakan SMOTE untuk mengatasi masalah ketika kita tidak memiliki data yang seimbang, pada penelitian ini kelas netral paling rendah performanya karena sulit dibedakan oleh model antara ulasan positif atau negatif. Selain itu, penelitian selanjutnya dapat menggunakan model IndoBERT untuk memahami makna ulasan teks bahasa Indonesia yang menggunakan singkatan atau bahasa tidak baku.

Daftar Rujukan

- [1] D. A. W. P., J. D. Firizqi, and Z. A. Amalia, 2024. Analisi Sentiment del prodotto skincare Somethinc Niacinamide su Female Daily con Naïve Bayes Classifier. *Journal Media Informatika Budidarma*, vol. 8, no. 2, pp. 946–956.
- [2] R. C. Larasati, C. Dewi, and H. J. Christanto, 2024. Analisis Sentimen Produk Kecantikan Jenis Moisturizer di Twitter Menggunakan Algoritma Support Vector Machine. *Journal TEKINKOM*, vol. 7, no. 1, pp. 1–12.

- [3] “Analisis Sentimen Ulasan Produk Skincare Daviena di Shopee Menggunakan Algoritma Support Vector Machine,” 2024. Seminar Nasional AMIKOM Surakarta.
- [4] E. Indrayuni, 2019. Klasifikasi Text Mining Review Produk Kosmetik untuk Teks Bahasa Indonesia Menggunakan Algoritma Naive Bayes. *Journal Khatulistiwa Informatika*, vol. 7, no. 1, pp. 29–36.
- [5] J. A. Putra, A. Dharmawan, and J. Gondohanindijo, 2024. Analisis Sentimen Aplikasi Digitalent Mobile Menggunakan Naive Bayes dan SVM dengan Ekstraksi Fitur TF-IDF. *Journal of Information Technology and Computer Science (INTECOMS)*, vol. 7, no. 4, pp. 1139–1148.
- [6] R. Vincent, I. Maulana, and O. Komarudin, 2023. Perbandingan Klasifikasi Naive Bayes dan Support Vector Machine dalam Analisis Sentimen dengan Multiclass di Twitter. *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 7, no. 4, pp. 2496–2505.
- [7] W. Ningsih, B. Alfianda, Rahmadden, and D. Wulandari, 2024. Perbandingan Algoritma SVM dan Naive Bayes dalam Analisis Sentimen Twitter pada Penggunaan Mobil Listrik di Indonesia. *MALCOLM: Indonesian Journal of Machine Learning and Computer Science*, vol. 4, no. 2, pp. 556–562.
- [8] J. A. Putra, A. Dharmawan, and J. Gondohanindijo, 2024. Analisis Sentimen Aplikasi Digitalent Mobile Menggunakan Naive Bayes dan SVM dengan Ekstraksi Fitur TF-IDF. *Journal of Information Technology and Computer Science (INTECOMS)*, vol. 7, no. 4, pp. 1139–1148.
- [9] “Analisis Sentimen Ulasan E Commerce Menggunakan Naive Bayes dan Support Vector Machine,” 2023. *Journal Infotech*.
- [10] Muszaki, V. Febriana, and W. Cholifah, 2024. Analisis Sentimen pada Ulasan Produk di E-Commerce Menggunakan Metode Naive Bayes. *Jurnal Riset dan Aplikasi Mahasiswa Informatika (JERAMI)*, vol. 5, no. 4, pp. 1–10.
- [11] “Analisis Sentimen Produk pada E-Commerce Menggunakan Naive Bayes dengan Akurasi Tinggi,” 2023. *Jurnal Teknologi Informasi*.
- [12] “Penerapan Support Vector Machine untuk Analisis Sentimen Data Teks,” 2022. *Jurnal Ilmu Komputer*.
- [13] Kurniawan, A. Hananto, S. Hilabi, et al., 2023. Perbandingan Algoritma Naive Bayes dan SVM dalam Sentimen Analisis Marketplace pada Twitter. *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, vol. 10, no. 1, pp. 731–740.
- [14] “Perbandingan Naive Bayes dan SVM dengan SMOTE pada Analisis Sentimen,” 2024. *Jurnal Teknologi dan Sistem Informasi*.
- [15] “Penerapan SVM dan TF-IDF untuk Analisis Sentimen Ulasan Aplikasi,” 2023. *Journal Teknologi Komputer*.
- [16] E. Indrayuni, 2019. Klasifikasi Text Mining Review Produk Kosmetik untuk Teks Bahasa Indonesia Menggunakan Algoritma Naive Bayes. *Journal Khatulistiwa Informatika*, vol. 7, no. 1, pp. 29–36.
- [17] Widyananda et al., 2025. Perbandingan Kinerja Model IndoBERT, IndoBERTweet, dan SVM pada Analisis Sentimen Ulasan E-commerce. *Journal Seminar ID*, vol. 1, no. 1, pp. 1–10.
- [18] H. Jayadianti et al., 2022. Sentiment Analysis of Indonesian Reviews Using Fine-Tuning IndoBERT and R CNN. *ILKOM Jurnal Ilmiah*, vol. 14, no. 3, pp. 348–354.
- [19] Nasution et al., 2025. Transformasi Pembelajaran Bahasa Indonesia melalui Deep Learning dan NLP. *Eunoia: Jurnal Pemikiran dan Penelitian Pendidikan Bahasa dan Seni*, vol. 5, no. 1, pp. 1–15.
- [20] H. N. Muchammad Shiddieqy, S. P. Insap, and W. W. Wahyu, 2016. Studi Literatur Tentang Perbandingan Metode untuk Proses Analisis Sentimen di Twitter. *Seminar Nasional Teknologi Informasi dan Komunikasi*, vol. 7, no. 2, pp. 57–64.
- [21] Natuzzuhriyyah, N. Nafisah, and R. Mayasari, 2021. Klasifikasi Tingkat Kepuasan Mahasiswa terhadap Pembelajaran Secara Daring Menggunakan Algoritma Naive Bayes. *JISKA*, vol. 6, no. 3, pp. 161–170.

- [22] H. Apriyani dan Kurniati, "Perbandingan Metode Naïve Bayes dan Support Vector Machine dalam Klasifikasi Penyakit Diabetes Melitus," *Journal of Information Technology Ampera*, vol. 1, no. 3, Des. 2020.