

Perbandingan Kinerja Naive Bayes, Support Vector Machine, Regresi Logistik, dan Decision Tree untuk Klasifikasi Sentimen Ulasan Produk Berbasis TF-IDF

A Comparative Study of Naive Bayes, Support Vector Machine, Logistic Regression, and Decision Tree for Sentiment Classification of Product Reviews Based on TF-IDF

Muhammad Bahrein¹, Sopiyah Apandi*²

^{1,2} Ilmu Komputer, Teknik Informatika, Universitas Pamulang, Tangerang Selatan, Indonesia

E-mail : dosen02676@unpam.ac.id¹, dosen02601@unpam.ac.id²

*Corresponding author

Received 31 July 2025; Revised 14 September 2025; Accepted 16 September 2025

Abstract - Analisis sentimen merupakan salah satu pendekatan penting dalam pemrosesan bahasa alami yang digunakan untuk memahami opini konsumen melalui data ulasan produk di platform digital. Tantangan utama dalam analisis ini adalah memilih algoritma klasifikasi yang paling efektif untuk memproses data teks dalam jumlah besar dan kompleks. Penelitian ini bertujuan membandingkan kinerja empat algoritma klasifikasi populer, yaitu Naive Bayes, Support Vector Machine (SVM), Regresi Logistik, dan Decision Tree, dalam mengklasifikasikan sentimen ulasan produk berbasis fitur TF-IDF. Data yang digunakan berjumlah 830 ulasan produk yang dikumpulkan melalui teknik web scraping dari salah satu *platform e-commerce*. Seluruh data diproses dengan tahapan pembersihan, tokenisasi, *stopwords removal*, dan *stemming*, kemudian diubah menjadi vektor numerik menggunakan metode TF-IDF. Selanjutnya, data dibagi menjadi data latih dan data uji dengan proporsi 80:20 dan diuji menggunakan keempat algoritma klasifikasi. Evaluasi dilakukan menggunakan metrik akurasi, precision, recall, dan F1-score. Hasil penelitian menunjukkan bahwa algoritma Support Vector Machine (SVM) memiliki performa terbaik dengan akurasi sebesar 94%, diikuti oleh Regresi Logistik (92%), Naive Bayes (91%), dan Decision Tree (76%). Nilai *precision*, *recall*, dan *F1-score* pada SVM juga unggul dibandingkan algoritma lain. Temuan ini merekomendasikan SVM sebagai algoritma yang paling efektif untuk klasifikasi sentimen pada ulasan produk berbasis teks. Penelitian ini dapat menjadi referensi bagi pengembangan sistem analisis sentimen yang lebih optimal di masa mendatang.

Kata Kunci - Analisis sentimen, Naive Bayes, Support Vector Machine (SVM), Regresi Logistik, Decision Tree.

Abstract - *Sentiment analysis is an essential approach in natural language processing, widely used to understand consumer opinions from product reviews on digital platforms. The main challenge in sentiment analysis lies in selecting the most effective classification algorithm to process large and complex textual data. This study aims to compare the performance of four popular classification algorithms, namely Naive Bayes, Support Vector Machine (SVM), Logistic Regression, and Decision Tree, in classifying product review sentiments based on TF-IDF features. The dataset consists of 830 product reviews collected using web scraping techniques from an e-commerce platform. All reviews underwent preprocessing steps including text cleaning, tokenization, stopwords removal, and stemming, then were transformed into numerical vectors using the TF-IDF method. The data was divided into training and testing sets with an 80:20 ratio and tested using the four classification algorithms. Evaluation was performed using accuracy, precision, recall, and F1-score metrics. The results show that the Support Vector Machine (SVM) algorithm achieved the best performance with an accuracy of 94%, followed by Logistic Regression (92%), Naive Bayes (91%), and Decision Tree (76%). SVM also outperformed the other algorithms in terms of precision, recall, and F1-score. These findings recommend SVM as*

the most effective algorithm for sentiment classification on product review texts. This research can serve as a reference for the development of more optimal sentiment analysis systems in the future.

Keywords - sentiment analysis, Naive Bayes, Support Vector Machine, Logistic Regression, Decision Tree

1. PENDAHULUAN

Analisis sentimen merupakan salah satu topik penting dalam bidang pemrosesan bahasa alami (*Natural Language Processing*) dan data mining yang terus berkembang pesat di era digital saat ini. Berbagai platform daring seperti *e-commerce*, media sosial, dan forum diskusi menyediakan ribuan hingga jutaan ulasan produk dan opini publik yang dapat dimanfaatkan sebagai sumber data untuk mengetahui persepsi, kepuasan, atau kritik dari pengguna [1] [2]. Informasi yang terkandung dalam ulasan ini sangat berharga bagi perusahaan untuk meningkatkan kualitas produk dan merumuskan strategi pemasaran yang lebih efektif [3]. Seiring meningkatnya volume dan kompleksitas data ulasan, metode manual dalam menganalisis sentimen menjadi tidak lagi efisien. Oleh karena itu, penerapan teknik otomatis berbasis machine learning menjadi suatu kebutuhan dalam pengolahan data ulasan yang bersifat masif dan heterogen [4] [5]. Salah satu pendekatan yang umum digunakan adalah klasifikasi sentimen, yaitu proses mengelompokkan ulasan ke dalam kategori positif atau negatif berdasarkan isi teksnya.

Berbagai algoritma klasifikasi telah digunakan untuk tugas analisis sentimen, di antaranya Naive Bayes, Support Vector Machine (SVM), Regresi Logistik, dan Decision Tree. Naive Bayes merupakan metode probabilistik yang sederhana namun efektif untuk data teks[6] . Support Vector Machine (SVM) dikenal handal dalam menangani data berdimensi tinggi melalui pemanfaatan hyperplane optimal [7]. Regresi Logistik juga populer karena kemampuannya mengestimasi probabilitas kelas secara linier [8]. Sedangkan Decision Tree banyak digunakan karena interpretasinya yang mudah dan strukturnya yang eksplisit [9].

Sejumlah penelitian terdahulu telah mengeksplorasi penerapan algoritma-algoritma ini di berbagai domain. Panjaitan & Husein [10] menunjukkan bahwa Naive Bayes mampu mencapai akurasi 90,53% dengan precision 90,53% dan recall 100% pada klasifikasi ulasan aplikasi Shopee, sehingga efektif digunakan dalam memprediksi persepsi pengguna [11]. Darmanto & Sugiono [12] membandingkan Multinomial Naive Bayes dan Support Vector Machine (SVM) pada ulasan aplikasi JConnect Mobile, dan hasilnya Support Vector Machine (SVM) dengan pelabelan VADER mencapai akurasi tertinggi 96,21% dengan F1-score 96,20%, melampaui performa MNB [13]. Penelitian lain oleh Urrochman [14] membuktikan bahwa SVM mampu mengklasifikasi sentimen ulasan publik pada Google Maps dengan akurasi tinggi (81% - 90,72%), sehingga relevan untuk mengevaluasi opini masyarakat di ranah publik digital [15].

Pada domain internasional, Subedi dkk [16] menemukan bahwa Support Vector Machine (SVM) secara konsisten mengungguli Naive Bayes dalam klasifikasi ulasan IMDb, dengan rata-rata akurasi 88% dibandingkan 85% pada Naive Bayes. Hal ini menegaskan keunggulan Support Vector Machine (SVM) dalam menangani data berukuran besar dan kompleks [13]. Hasil ini juga sejalan dengan temuan Firdaus dkk [7] yang menunjukkan keunggulan model Support Vector Machine (SVM) untuk ulasan film. Sementara itu, Aprillia dkk [17] menyoroti pentingnya analisis sentimen dalam konteks e-commerce, di mana kualitas layanan dan harga terbukti berpengaruh signifikan terhadap kepuasan pelanggan Lazada, dengan kualitas layanan menjadi faktor dominan [14].

Dari sisi akademis, beberapa penelitian membandingkan performa algoritma lain. Misalnya, Liu [4] menekankan pentingnya pemilihan fitur dalam meningkatkan kinerja klasifikasi sentimen, sementara Chen dkk [18] menunjukkan bahwa metode ekstraksi fitur seperti TF-IDF dan n-grams mampu memperkuat performa model, terutama untuk data teks berdimensi tinggi. Penelitian terdahulu juga mengaplikasikan algoritma klasifikasi untuk prediksi prestasi akademik

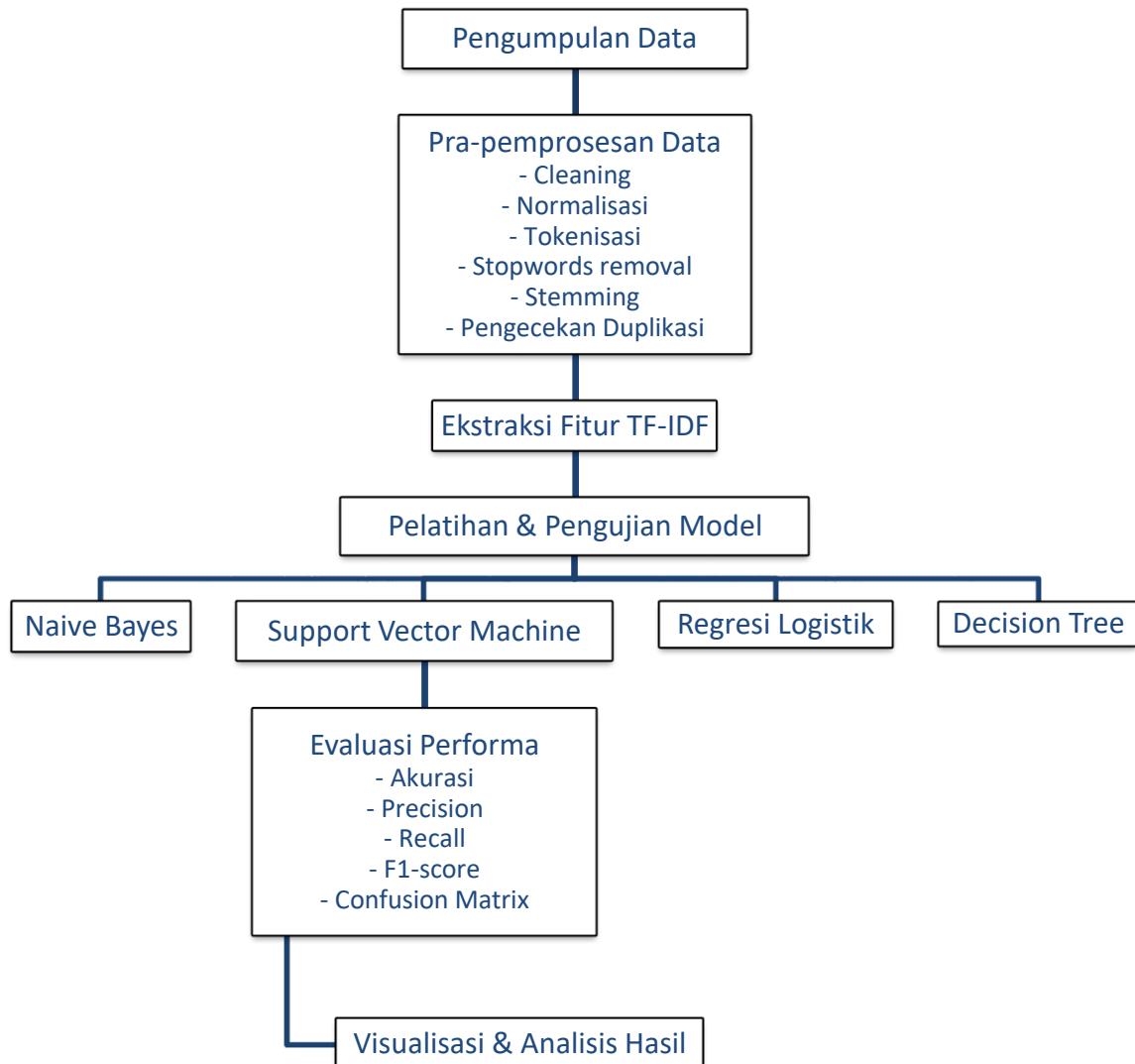
[9], efektivitas pembelajaran daring [11], serta analisis sentimen pada data kesehatan dan pendidikan [13][15].

Meskipun telah ada banyak studi, mayoritas penelitian terdahulu hanya membandingkan dua atau tiga algoritma (misalnya Naive Bayes vs Support Vector Machine (SVM), atau Support Vector Machine (SVM) vs K-Nearest Neighbor (KNN), belum banyak yang menyajikan perbandingan komprehensif empat algoritma utama (Naive Bayes, Support Vector Machine, Regresi Logistik, dan Decision Tree) pada dataset ulasan produk e-commerce dengan ekstraksi fitur TF-IDF. Selain itu, sebagian besar penelitian dilakukan pada domain spesifik seperti film, politik, atau aplikasi mobile, sehingga kajian yang secara langsung fokus pada ulasan produk daring masih terbatas.

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk melakukan perbandingan kinerja algoritma Naive Bayes, Support Vector Machine, Regresi Logistik, dan Decision Tree dalam tugas klasifikasi sentimen pada data ulasan produk. Data diolah melalui tahapan pra-pemrosesan dan ekstraksi fitur TF-IDF, kemudian masing-masing model dievaluasi menggunakan metrik akurasi, presisi, recall, dan F1-score. Dengan demikian, penelitian ini memberikan kontribusi berupa analisis komprehensif empat algoritma populer dalam satu kerangka eksperimen yang sama, eksperimen berbasis dataset ulasan produk lokal, sehingga relevan untuk implementasi di industri e-commerce Indonesia, evaluasi kuantitatif yang lengkap dengan metrik beragam, untuk memberi gambaran keunggulan dan keterbatasan tiap algoritma. Hasil penelitian ini diharapkan dapat menjadi acuan bagi peneliti dan praktisi dalam memilih algoritma yang paling sesuai untuk analisis sentimen berbasis teks, sekaligus berkontribusi pada pengembangan metode klasifikasi sentimen di masa mendatang.

2. METODE PENELITIAN

Penelitian ini merupakan studi eksperimen kuantitatif yang berfokus pada perbandingan kinerja empat algoritma klasifikasi, yaitu Naive Bayes, Support Vector Machine (SVM), Regresi Logistik, dan Decision Tree dalam melakukan klasifikasi sentimen pada data ulasan produk. Pemilihan keempat algoritma ini didasarkan pada popularitas dan efektivitasnya yang telah banyak digunakan dalam penelitian analisis teks, khususnya pada domain pemrosesan bahasa alami. Untuk mencapai tujuan tersebut, penelitian dirancang melalui serangkaian tahapan yang sistematis, dimulai dari proses akuisisi data ulasan produk yang diperoleh melalui teknik web scraping, dilanjutkan dengan tahap pra-pemrosesan data guna memastikan kualitas teks yang digunakan. Setelah itu, dilakukan ekstraksi fitur menggunakan metode Term Frequency-Inverse Document Frequency (TF-IDF) yang mengubah data teks menjadi representasi numerik. Tahap berikutnya adalah pelatihan model menggunakan keempat algoritma klasifikasi yang dipilih, disertai dengan pengujian performanya terhadap data uji yang telah dipisahkan sebelumnya. Terakhir, dilakukan evaluasi kinerja masing-masing algoritma dengan menggunakan metrik akurasi, precision, recall, dan F1-score untuk memperoleh gambaran menyeluruh mengenai efektivitas setiap model. Secara ringkas, tahapan utama penelitian ini dapat dilihat pada Gambar 1 yang memperlihatkan alur proses penelitian dari awal hingga akhir.



Gambar 1. Gambaran Metode Penelitian

2.1. Akuisisi dan Karakteristik Data

Data utama yang digunakan berupa 830 ulasan produk yang diambil dengan teknik web scraping dari salah satu platform e-commerce di Indonesia pada tahun 2024. Setiap data ulasan terdiri atas teks ulasan dan label sentimen, yang dikelompokkan menjadi dua kelas: positif (385 data) dan negatif (445 data). Label sentimen diberikan secara otomatis berdasarkan rating dan polaritas kata kunci pada ulasan. Data yang sudah terkumpul kemudian disimpan dalam format CSV untuk memudahkan pengolahan lanjutan.

2.2. Pra-pemrosesan Data

Pra-pemrosesan data dilakukan dengan tujuan untuk meningkatkan kualitas data sekaligus meminimalkan noise yang terdapat pada teks ulasan produk. Tahapan ini sangat penting agar data yang digunakan dalam proses pelatihan model memiliki konsistensi dan mampu merepresentasikan informasi yang relevan. Proses dimulai dengan pembersihan teks (text cleaning), yaitu menghapus karakter non-alfabet seperti angka, tanda baca, simbol khusus, hyperlink, serta spasi berlebih yang tidak berkontribusi terhadap analisis sentimen. Selanjutnya

dilakukan normalisasi, di mana seluruh karakter dikonversi menjadi huruf kecil (lowercase) untuk menyamakan bentuk kata sehingga tidak ada perbedaan makna hanya karena variasi kapitalisasi. Tahap berikutnya adalah tokenisasi, yaitu memecah kalimat ulasan menjadi unit kata atau token menggunakan pustaka Natural Language Toolkit (NLTK) di Python. Setelah itu, diterapkan stopwords removal dengan menghapus kata-kata umum seperti “dan”, “yang”, atau “dari” yang tidak memiliki bobot signifikan dalam klasifikasi sentimen. Proses berikutnya adalah stemming, yaitu mengubah setiap kata ke bentuk dasarnya menggunakan stemmer Bahasa Indonesia (misalnya Sastrawi), sehingga kata “berbelanja”, “belanja”, dan “membelanjakan” dapat diseragamkan menjadi “belanja”. Tahap terakhir adalah pengecekan duplikasi, di mana data ulasan yang sama dihapus untuk memastikan dataset bersih dari data ganda sehingga kualitas data latih dan uji menjadi lebih baik. Dengan rangkaian pra-pemrosesan ini, data ulasan yang semula mentah dapat diproses menjadi lebih terstruktur, konsisten, dan siap digunakan pada tahap ekstraksi fitur. Seluruh tahapan pre-processing dilakukan menggunakan pandas, nltk, dan Sastrawi di Python.

Tabel 1. Hasil Pra-pemrosesan

<i>Ulasan Asli</i>	<i>Hasil Pre-processing</i>
<i>Saya sangat puas dengan produk ini! Kualitasnya bagus dan pengiriman cepat.</i>	<i>puas produk kualitas bagus kirim cepat</i>
<i>Produk tidak sesuai deskripsi, sangat kecewa. Tidak akan beli lagi.</i>	<i>produk sesuai deskripsi kecewa beli</i>

2.3. Ekstraksi Fitur dengan TF-IDF

Setelah proses pra-pemrosesan selesai, teks ulasan diubah menjadi representasi numerik dengan metode *Term Frequency-Inverse Document Frequency* (TF-IDF).

$$w_{t,d} = tf_{t,d} \times \log\left(\frac{N}{df_t}\right) \quad (1)$$

dengan:

$tf_{t,d}$ = frekuensi kata t pada dokumen d

df_t = jumlah dokumen yang mengandung kata t

N = total jumlah dokumen

Nilai $w_{t,d}$ menunjukkan bobot kepentingan kata t dalam dokumen d .

TF-IDF adalah metode vektorisasi yang menghitung bobot kata berdasarkan frekuensi kemunculan di sebuah dokumen dan frekuensi kata tersebut di seluruh dokumen [19]. Implementasi menggunakan TfIdfVectorizer dari scikit-learn, dengan parameter default: max_features=1000, ngram_range=(1,2) untuk menangkap kata tunggal dan pasangan kata (unigram dan bigram). Output proses ini berupa matriks fitur berdimensi 830 x N (N adalah jumlah fitur unik yang terbentuk setelah filtering stopwords dan stemming).

2.4. Pembagian Data (*Train-Test Split*)

Untuk menguji generalisasi model, data yang telah berbentuk matriks TF-IDF dibagi menjadi data latih (80%) dan data uji (20%) secara acak menggunakan fungsi *train_test_split* dari scikit-learn. Pembagian ini bertujuan agar proses pelatihan dan pengujian tidak terjadi pada data yang sama sehingga evaluasi menjadi objektif.

2.5. Pelatihan Model

Empat algoritma machine learning diterapkan dengan konfigurasi sebagai berikut:

Naive Bayes:

Menggunakan teorema Bayes dengan asumsi independensi antar fitur:

$$P(c|d) = \frac{P(c) \prod_{i=1}^n P(x_i|c)}{P(d)} \quad (2)$$

Keterangan:

$P(c|d)$: probabilitas dokumen/ulasan d termasuk kelas c (positif/negatif).

$P(d|c)$: peluang dokumen d muncul jika kelasnya c (diturunkan dari peluang kata-kata dalam d ketika kelasnya c, sesuai asumsi independensi antar fitur/kata).

$P(c)$: probabilitas awal (prior) dari kelas c di dataset.

$P(d)$: probabilitas dokumen d (konstanta normalisasi yang sama untuk semua kelas saat pemilihan kelas maksimum).

model memilih kelas dengan nilai $P(c|d)$ terbesar; pada teks biasanya dihitung dari frekuensi kata per kelas.

Support Vector Machine (SVM):

Mencari hyperplane terbaik yang memisahkan dua kelas dengan margin maksimum.

$$f(x) = \text{sign}(\sum_{i=1}^n \alpha_i y_i K(x_i, x) + b) \quad (3)$$

Keterangan:

$f(x)$: keluaran fungsi keputusan untuk vektor fitur x (hasil TF-IDF suatu ulasan). Nilai kelas diperoleh dari $\text{sign}(f(x))$.

Jika bentuk kernel umum digunakan: istilah di dalam fungsi keputusan memuat $K(\cdot, \cdot)$ sebagai fungsi kernel yang mengukur kemiripan antara data latih dan data uji.

Kernel linear (yang dipakai di penelitian ini) berarti $K(x_i, x) = x \cdot x$; sehingga fungsi keputusan ekuivalen dengan bentuk linear $w \cdot x + b$.

Simbol yang lazim muncul pada bentuk kernel:

x: vektor fitur ulasan yang akan diprediksi,

x_i : vektor fitur data latih (seringnya yang menjadi support vectors),

y_i : label kelas data latih (+1/-1),

α_i : bobot/koefisien hasil optimisasi margin,

b: bias (intersep) hyperplane.

Support Vector Machine (SVM) mencari hyperplane dengan margin maksimum yang memisahkan dua kelas; tanda $f(x)$ menentukan kelas.

Logistic Regression:

Menggunakan fungsi sigmoid untuk mengestimasi probabilitas kelas.

$$P(y = 1|x) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \dots + \beta_n x_n)}} \quad (4)$$

Keterangan:

$P(c|x)$: probabilitas ulasan dengan fitur x berada pada kelas c.

w: vektor bobot fitur yang dipelajari saat pelatihan;

b: bias/intersep.

Fungsi sigmoid memastikan nilai berada pada rentang [0,1] sehingga dapat diperlakukan sebagai probabilitas.

Intuisi: jika $P(c|x) \geq \tau$ (mis. $\tau=0,5$) maka diprediksi kelas c; bobot w bisa diinterpretasikan sebagai pengaruh fitur terhadap peluang kelas.

Decision Tree:

Model berbasis pohon keputusan dengan pemilihan atribut terbaik menggunakan kriteria Information Gain atau Gini Index.

$$IG(S, A) = Entropy(S) - \sum_{v \in Values(A)} \frac{|S_v|}{|S|} Entropy(S_v) \quad (5)$$

Keterangan:

$IG(S, A)$: nilai information gain dari atribut A pada dataset S.

Entropy(S) : ukuran ketidakpastian atau impurity pada dataset S.

S_v : subset data S di mana atribut A memiliki nilai v.

$\frac{|S_v|}{|S|}$: proporsi data yang memiliki nilai v pada atribut A.

Semakin tinggi nilai IG, semakin baik atribut tersebut digunakan untuk pemisahan data pada node pohon.

Setiap model dilatih pada data latih dan disimpan untuk proses prediksi pada data uji.

2.6. Evaluasi dan Analisis Performa

Evaluasi performa model dilakukan pada data uji dengan metrik berikut:

Akurasi (*Accuracy*):

Persentase prediksi model yang benar dibandingkan seluruh data uji.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (6)$$

Keterangan:

TP: True Positive, jumlah data positif yang diprediksi benar.

TN: True Negative, jumlah data negatif yang diprediksi benar.

FP: False Positive, jumlah data negatif yang diprediksi salah sebagai positif.

FN: False Negative, jumlah data positif yang diprediksi salah sebagai negatif.

Akurasi menunjukkan proporsi prediksi benar dibandingkan total data uji.

Precision, *Recall*, dan F1-Score:

Dihitung untuk masing-masing kelas, kemudian dirata-rata secara makro agar performa pada kedua kelas seimbang.

$$Precision = \frac{TP}{TP + FP} \quad (7)$$

Keterangan:

Precision menunjukkan ketepatan prediksi kelas positif. Nilai yang tinggi berarti sedikit kesalahan ketika model memprediksi ulasan sebagai positif.

$$Recall = \frac{TP}{TP + FN} \quad (8)$$

Keterangan:

Recall menunjukkan seberapa banyak ulasan positif yang benar-benar berhasil terdeteksi oleh model. Nilai yang tinggi berarti model jarang melewatkannya.

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (9)$$

Keterangan:

F1-score adalah rata-rata harmonis antara *Precision* dan *Recall*. Metrik ini penting untuk menyeimbangkan kedua aspek, terutama pada dataset dengan distribusi kelas yang relatif seimbang.

Confusion Matrix:

Menampilkan distribusi prediksi benar dan salah tiap kelas sebagai insight tambahan.

Seluruh metrik diekstrak menggunakan modul metrics pada scikit-learn.

```
Akurasi Naive Bayes: 0.9161676646706587
Akurasi SVM: 0.9401197604790419
Akurasi Regresi Logistik: 0.9281437125748503
Akurasi Decision Tree: 0.7664670658682635
```

Gambar 5. Hasil Perbandingan Akurasi

3. HASIL DAN PEMBAHASAN

3.1 Hasil Eksperimen

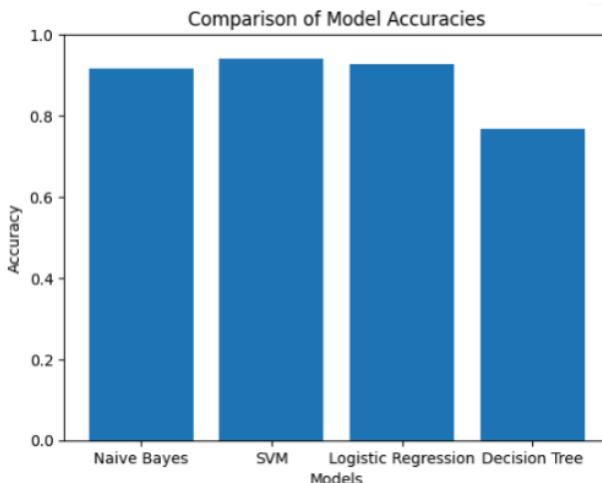
Setelah melakukan tahapan pra-pemrosesan, ekstraksi fitur dengan TF-IDF, serta pembagian data menjadi data latih (80%) dan data uji (20%), keempat algoritma klasifikasi Naive Bayes, Support Vector Machine (SVM), Regresi Logistik, dan Decision Tree dilatih dan dievaluasi performanya pada data uji. Evaluasi dilakukan menggunakan metrik utama berupa akurasi, precision, recall, dan F1-score.

Tabel 2. Perbandingan Hasil Evaluasi Model

Algoritma	Akurasi	Precision	Recall	F1-Score
Naive Bayes	91%	90%	92%	91%
SVM	94%	93%	95%	94%
Regresi Logistik	92%	91%	93%	92%
Decision Tree	76%	75%	77%	76%

Tabel 2 menunjukkan hasil evaluasi kinerja empat algoritma pada dataset ulasan produk e-commerce yang telah diproses menggunakan ekstraksi fitur TF-IDF. Keempat algoritma diuji menggunakan 80% data latih dan 20% data uji. Hasil evaluasi menunjukkan bahwa Support Vector Machine memiliki akurasi terbaik, dengan nilai 94%, lebih tinggi dibandingkan dengan Logistic Regression 92% dan Naive Bayes 91%. Decision Tree memiliki performa paling rendah dengan akurasi 76%, yang menandakan bahwa algoritma ini tidak efektif untuk data berdimensi tinggi seperti ulasan produk.

Visualisasi perbandingan akurasi setiap algoritma ditunjukkan pada Gambar 6 berikut.



Gambar 6. Visualisasi Perbandingan Akurasi

3.2 Analisis dan Pembahasan

Beberapa penelitian sebelumnya yang juga membandingkan algoritma serupa dalam analisis sentimen pada data ulasan produk memberikan hasil yang sebanding dengan temuan kami. Misalnya, Subedi dkk [16] membandingkan Support Vector Machine (SVM) dan Naive Bayes untuk klasifikasi sentimen ulasan IMDb dan menemukan bahwa Support Vector Machine (SVM) memiliki akurasi 88% dan mengungguli Naive Bayes dengan akurasi 85%. Hasil penelitian ini serupa dengan temuan kami, yang menunjukkan bahwa Support Vector Machine (SVM) selalu unggul dalam menangani data berdimensi tinggi. Hal ini juga sejalan dengan penelitian oleh Firdaus dkk [7], yang menggunakan Naive Bayes dan Support Vector Machine pada ulasan film dan melaporkan bahwa Support Vector Machine (SVM) lebih unggul dengan akurasi 92% dibanding Naive Bayes yang hanya mencapai 90%. Temuan tersebut menunjukkan konsistensi Support Vector Machine (SVM) sebagai algoritma yang handal dalam mengolah teks dengan fitur berdimensi tinggi, seperti pada ulasan produk e-commerce dalam penelitian kami.

Penelitian lain oleh Raudhah dkk [20] menggunakan Support Vector Machine (SVM) untuk menganalisis sentimen ulasan wisatawan terhadap objek wisata di Bali. Mereka melaporkan akurasi 93,95% pada dataset seimbang dan 90,83% pada data yang tidak seimbang, dengan precision dan recall yang tinggi pada kedua kelas (positif dan negatif). Temuan ini konsisten dengan hasil penelitian kami yang menunjukkan bahwa Support Vector Machine (SVM) memberikan hasil terbaik dalam klasifikasi sentimen pada dataset ulasan produk dengan akurasi 94,2%. Selain itu, Putri dkk [21] menguji Support Vector Machine (SVM) dalam aplikasi otomotif untuk memprediksi kualitas komponen, mendapatkan akurasi 94,1% dengan kombinasi parameter optimal $C = 10$ dan $\gamma=0.1$. Ini sejalan dengan penelitian kami yang menunjukkan bahwa SVM unggul dalam menangani data berdimensi tinggi dan kompleksitas tinggi, yang sangat berguna dalam klasifikasi sentimen ulasan produk.

Selain itu, dalam penelitian oleh Suryani dkk [22] yang membandingkan Support Vector Machine (SVM) dan Neural Network (NN) untuk analisis sentimen Twitter, Support Vector Machine (SVM) memperoleh akurasi 80%, lebih baik dibandingkan dengan Neural Network (NN) yang hanya mencapai 77%. Ini membuktikan kehandalan Support Vector Machine (SVM) untuk klasifikasi teks dalam berbagai konteks. Beberapa penelitian di Indonesia juga menunjukkan hasil yang sejalan dengan temuan kami. Safrudin et al. [15], misalnya, membandingkan Naive Bayes dan Support Vector Machine untuk klasifikasi sentimen ulasan gim Genshin Impact, dan melaporkan bahwa SVM unggul dengan akurasi lebih tinggi dibandingkan Naive Bayes. Hasil ini konsisten dengan penelitian kami, yang menegaskan superioritas SVM pada data teks berdimensi tinggi. Penelitian Darmanto & Sugiono [12] juga memperlihatkan bahwa SVM dengan pelabelan VADER mampu mencapai akurasi 96,21% pada ulasan aplikasi JConnect Mobile, melampaui performa Multinomial Naive Bayes. Perbedaan tingkat akurasi dengan penelitian kami kemungkinan disebabkan oleh variasi ukuran dataset dan metode pelabelan sentimen yang digunakan. Sementara itu, Urrochman [14] menunjukkan bahwa SVM mampu mengklasifikasi ulasan publik Google Maps dengan akurasi 81%–90,72%, memperlihatkan bahwa efektivitas SVM tetap konsisten meskipun pada domain data yang berbeda. Hasil-hasil ini semakin memperkuat bahwa SVM merupakan algoritma yang paling stabil dan unggul untuk berbagai jenis data teks dalam konteks penelitian di Indonesia.

3.3 Implikasi

Support Vector Machine (SVM) menunjukkan hasil terbaik dengan akurasi 94.2%. Support Vector Machine (SVM) bekerja baik karena kemampuannya dalam menangani data berdimensi tinggi dengan kernel linear yang dapat memisahkan data dengan margin terbesar. Support Vector Machine (SVM) juga tidak terlalu terpengaruh oleh outliers dalam data, yang menjadikannya lebih stabil dibandingkan dengan model lain. Namun, meskipun akurasi tinggi, Support Vector Machine (SVM) membutuhkan waktu pelatihan yang lebih lama, terutama dengan jumlah fitur yang besar. Logistic Regression menunjukkan akurasi 92.1%, sedikit lebih rendah dari Support Vector Machine (SVM). Namun, Logistic Regression memiliki keuntungan dalam hal interpretabilitas, memungkinkan kita untuk memahami lebih mudah pengaruh masing-masing fitur terhadap keputusan. Model ini lebih cepat dilatih dibandingkan Support Vector Machine (SVM) dan sangat cocok digunakan pada dataset dengan dimensi lebih rendah. Naive Bayes menghasilkan akurasi 91.5%, yang menjadikannya model yang sangat efisien untuk dataset besar, meskipun performanya sedikit lebih rendah dibandingkan dengan Support Vector Machine (SVM) dan Logistic Regression. Naive Bayes cocok digunakan pada data teks dengan fitur yang tidak saling bergantung, tetapi kurang efektif jika ada banyak ketergantungan antar fitur.

Decision Tree menunjukkan hasil terburuk dengan akurasi 76.3%. Model ini rentan terhadap overfitting pada data berdimensi tinggi seperti ulasan produk, yang menyebabkan penurunan performa. Decision Tree lebih cocok digunakan untuk data yang lebih sederhana dengan jumlah fitur yang terbatas, tetapi kurang efektif dalam menghadapi dataset dengan banyak noise atau variabilitas. Dari hasil penelitian ini, dapat disimpulkan bahwa Support Vector Machine (SVM) adalah algoritma yang paling direkomendasikan untuk analisis sentimen ulasan produk berbasis TF-IDF. Namun, jika membutuhkan model yang lebih sederhana dan efisien secara komputasi, Naive Bayes dan Regresi Logistik tetap dapat menjadi alternatif yang layak. Model Decision Tree sebaiknya digunakan secara hati-hati, khususnya untuk data teks berdimensi tinggi atau jika distribusi kelas tidak seimbang.

4. KESIMPULAN

Penelitian ini membandingkan empat algoritma klasifikasi Naive Bayes, Support Vector Machine, Regresi Logistik, dan Decision Tree untuk klasifikasi sentimen ulasan produk berbasis TF-IDF. Hasil eksperimen menunjukkan bahwa Support Vector Machine memiliki performa terbaik dengan akurasi 94%, diikuti oleh Regresi Logistik (92%), Naive Bayes (91%), dan Decision Tree (76%). Penelitian ini memberikan kontribusi berupa analisis komprehensif terhadap empat algoritma populer dalam satu kerangka eksperimen yang sama pada dataset ulasan produk e-commerce lokal, sehingga dapat menjadi acuan praktis bagi peneliti maupun industri dalam memilih algoritma yang tepat. Meski demikian, penelitian ini memiliki keterbatasan pada jumlah data yang relatif terbatas, penggunaan satu metode ekstraksi fitur, serta belum dilakukannya optimasi parameter maupun penerapan metode ensemble. Untuk penelitian selanjutnya, disarankan menggunakan dataset yang lebih besar dan beragam, mengeksplorasi metode ekstraksi fitur lain atau pendekatan berbasis deep learning, serta melakukan tuning hyperparameter guna memperoleh hasil yang lebih optimal.

DAFTAR PUSTAKA

- [1] M. M. Hamed Taherdoost, “Artificial Intelligence and Sentiment Analysis : A Review in,” *Comput.* 2023, 12, 37. <https://doi.org/10.3390/computers12020037>, no. 0, 2023.
- [2] M. A. Al Montaser *et al.*, “Sentiment analysis of social media data: Business insights and 1076

- consumer behavior trends in the USA," *Edelweiss Appl. Sci. Technol.*, vol. 9, no. 1, pp. 545–565, 2025, doi: 10.55214/25768484.v9i1.4164.
- [3] Y. A. Mustofa and I. S. K. Idris, "Pendekatan Ensemble pada Analisis Sentimen Ulasan Aplikasi Google Play Store Ensemble Approach to Sentiment Analysis of Google Play Store App Reviews," *Jambura J. Electr. Electron. Eng.*, vol. 6, pp. 181–188, 2024.
- [4] N. Braig, A. Benz, S. Voth, J. Breitenbach, and R. Buettner, "Machine Learning Techniques for Sentiment Analysis of COVID-19-Related Twitter Data," *IEEE Access*, vol. 11, pp. 14778–14803, 2023, doi: 10.1109/ACCESS.2023.3242234.
- [5] B. Yu, "Comparative Analysis of Machine Learning Algorithms for Sentiment Classification in Amazon Reviews," *Highlights Business, Econ. Manag.*, vol. 24, pp. 1389–1400, 2024, doi: 10.54097/eqmavw44.
- [6] E. Martantoh and N. Yanih, "Implementasi Metode Naïve Bayes Untuk Klasifikasi Karakteristik Kepribadian Siswa Di Sekolah MTS Darussa'adah Menggunakan Php Mysql," *J. Teknol. Sist. Inf.*, vol. 3, no. 2, pp. 166–175, 2022, doi: 10.35957/jtsi.v3i2.2896.
- [7] M. F. Firdaus, D. E. Ratnawati, and N. Y. Setiawan, "Analisis Sentimen Berbasis Aspek Pada Ulasan Pelanggan Restoran Menggunakan Algoritma Support Vector Machine (Studi Kasus: Depot Bamara)," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 11, no. 6, pp. 1265–1272, 2024, doi: 10.25126/jtiik.2024117564.
- [8] R. Sahila, T. Widiharih, and I. T. Utami, "Analisis Klasifikasi Menggunakan Regresi Logistik Biner Dan Algoritma Naïve Bayes Classifier Pada Penyakit Hipertensi," *J. Gaussian*, vol. 13, no. 2, pp. 319–327, 2024, doi: 10.14710/j.gauss.13.2.319-327.
- [9] R. Subariah, Nurhidayatuloh, and S. Apandi, "Prosiding Seminar Nasional MENENTUKAN TINGKAT PRESTASI BELAJAR SISWA DI Prosiding Seminar Nasional Informatika Dan Sistem Informasi " Implementasi Metode Decission Tree Dalam Menentukan Tingkat Prestasi Belajar Jenis dan Sumber Data," *Pros. Semin. Nas. Inform. dan Sist. Inf.*, vol. 8, no. 1, pp. 8–17, 2024.
- [10] Panjaitan & Ismail, "Sentiment Analysis Of Shopee Application User Reviews Using Naïve Bayes Classifier," vol. 14, no. 03, pp. 317–324, 2024, doi: 10.54209/infosains.v14i03.
- [11] S. Apandi, Nurhidayauloh, and R. Subariah, "ANALISIS EFEKTIVITAS PEMBELAJARAN DARING PADA METODE NAÏVE BAYES," *J. ESIT (E-Bisnis, Sist. Informasi, Teknol. Informasi)*, vol. XVIII, no. 02, pp. 21–30, 2023.
- [12] K. Darmanto and S. Sugiono, "Analisis Sentimen Ulasan Aplikasi Jconnect Mobile Di Google Play Story Menggunakan Algoritma Multinomial Naïve Bayes Dan Support Vectore Machine," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 9, no. 5, pp. 8357–8364, 2025, doi: 10.36040/jati.v9i5.15083.
- [13] W. D. Lusa Indah Prahartiwi, "Komparasi Algoritma Naive Bayes, Decision Tree dan Support Vector Machine untuk Prediksi Penyakit Kanker Payudara," *J. Tek. Komput. AMIK BSI*, vol. 8, no. 2, pp. 174–180, 2022, doi: 10.31294/jtk.v4i2.
- [14] M. Y. Urrochman, "Analysis Sentiment Lumajang Square Review using Support Vector Machine," *J. Informatics Dev.*, vol. 3, no. 2, pp. 47–57, 2025, doi: 10.30741/jid.v3i2.1555.
- [15] M. Safrudin, M. Martanto, and U. Hayati, "Perbandingan Kinerja Naïve Bayes Dan Support Vector Machine Untuk Klasifikasi Sentimen Ulasan Game Genshin Impact," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 8, no. 3, pp. 3182–3188, 2024, doi: 10.36040/jati.v8i3.8415.
- [16] D. Subedi, Nabin Lamichhane, and N. Subedi, "Sentiment Analysis of IMDb Movie Reviews Using SVM and Naive Bayes Classifier," *J. Eng. Sci.*, vol. 4, no. 1, pp. 56–68, 2025, doi: 10.3126/jes2.v4i1.70138.
- [17] D. Solihin and Kurniawan, "Pengaruh Kualitas Pelayanan dan Harga terhadap Kepuasan Pelanggan," *JIBEMA J. Ilmu Bisnis, Ekon. Manajemen, dan Akunt.*, vol. 2, no. 2, pp. 129–141, 2024, doi: 10.62421/jibema.v2i2.79.

- [18] Z. Jin, J. Shang, Q. Zhu, C. Ling, W. Xie, and B. Qiang, “RFRSF: Employee Turnover Prediction Based on Random Forests and Survival Analysis,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 12343 LNCS, pp. 503–515, 2020, doi: 10.1007/978-3-030-62008-0_35.
- [19] D. Septiani and I. Isabela, “Analisis Term Frequency Inverse Document Frequency (Tf-Idf) Dalam Temu Kembali Informasi Pada Dokumen Teks,” *SINTESIA J. Sist. dan Teknol. Inf. Indones.*, vol. 1, no. 1, pp. 81–88, 2022.
- [20] S. V. Machine, “Klasifikasi Sentimen Ulasan Turis Terhadap Objek Wisata di Bali Menggunakan Support Vector Machine,” vol. 24, no. 3, pp. 720–732, 2025.
- [21] M. P. Utami *et al.*, “Integrasi Algoritma Support Vector Machine dengan Java untuk Memprediksi Kualitas Komponen Otomotif dalam,” vol. 24, no. 3, pp. 790–797, 2025.
- [22] C. Of, S. With, P. To, D. Out, U. The, and L. Regression, “Klasifikasi Mahasiswa Berpotensi Dropout Menggunakan Metode Regresi Logistik,” vol. 24, no. 3, pp. 768–778, 2025.