JURNAL ILMIAH FIFO
Vol. 17 No. 2 (2025) 212-224
DOI: 10.22441/fifo.2025.v17i2.010

P-ISSN 2085-4315
E-ISSN 2502-8332

Klasifikasi Sentimen iPhone Bekas di Tokopedia menggunakan Naïve Bayes dan Support Vector Machine

Krisna Novianto[1], Herlawati Herlawati[2*], Agus Hidayat[3]

[1,2,3] Informatika; Fakultas Ilmu Komputer; Universitas Bhayangkara Jakarta Raya
Email: [1]krisnanovianto2@gmail.com, [2] herlawati@ubharajaya.ac.id, [3]agus.hidayat@dsn.ubharajaya.ac.id
Penulis Korespondensi[*]

(***received*****:** 14-11-25, ***revised*****:** 22-11-25, ***accepted*****:** 25-11-25)

Abstrak

Kenaikan harga iPhone baru mendorong meningkatnya pembelian iPhone second di platform e-commerce seperti Tokopedia. Namun, konsumen masih menghadapi berbagai risiko terkait kondisi perangkat, performa komponen, dan keaslian yang umumnya teridentifikasi melalui ulasan pengguna. Penelitian ini bertujuan menganalisis sentimen dari 1.863 ulasan iPhone second untuk memperoleh gambaran objektif mengenai pengalaman konsumen. Teks ulasan diproses menggunakan TF-IDF sebagai representasi fitur dan SMOTE untuk mengatasi ketidakseimbangan kelas. Dua algoritma Naive Bayes dan Support Vector Machine (SVM) dibandingkan untuk menilai efektivitas klasifikasi. Hasil pengujian menunjukkan bahwa SVM memberikan performa terbaik dengan akurasi 96%, melampaui Naive Bayes yang mencapai 93%. Analisis lebih lanjut menemukan bahwa ulasan positif umumnya berkaitan dengan kualitas fisik dan kecepatan pengiriman, sedangkan ulasan negatif banyak menyoroti isu teknis serta keaslian perangkat. Penelitian ini berkontribusi pada penguatan literatur analisis sentimen e-commerce melalui evaluasi komprehensif terhadap kombinasi TF-IDF + SMOTE serta perbandingan performa Naive Bayes dan SVM dalam klasifikasi opini konsumen. Temuan ini menyediakan dasar empiris untuk penelitian lanjutan mengenai penilaian kualitas produk bekas berbasis ulasan daring.

Kata Kunci: Analisis Sentimen, iPhone Bekas, Tokopedia, Naive Bayes, Support Vector Machine

Abstract

The rising prices of new iPhone devices have driven many consumers to purchase second-hand units, particularly through e-commerce platforms such as Tokopedia. However, users still face various risks related to device condition, component performance, and authenticity—issues that are typically revealed through prior customer reviews. This study aims to analyze the sentiment of 1,863 reviews of second-hand iPhones to obtain an objective understanding of consumer experiences. The textual data were processed using TF-IDF for feature representation and SMOTE to address class imbalance. Two classification algorithms Naive Bayes and Support Vector Machine (SVM) were compared to evaluate their effectiveness. Experimental results indicate that SVM achieved the highest performance with 96% accuracy, surpassing Naive Bayes, which achieved 93% accuracy. Further analysis reveals that positive reviews commonly highlight physical quality and delivery speed, while negative reviews frequently address technical issues and device authenticity. This study contributes to the literature on e-commerce sentiment analysis by providing a comprehensive evaluation of the TF-IDF + SMOTE approach and a comparative assessment of Naive Bayes and SVM in classifying consumer opinions. The findings offer an empirical foundation for future research on quality assessment of second-hand products based on online reviews.

Keywords*: Sentiment Analysis, second-hand iPhone, Tokopedia, Naive Bayes, Support Vector Machine*

1. PENDAHULUAN

iPhone merupakan salah satu *smartphone* yang diminati oleh masyarakat Indonesia. Namun, harga seri terbaru iPhone yang mencapai puluhan juta rupiah sering dianggap terlalu tinggi oleh sebagian konsumen yang berasal dari kalangan menengah ke bawah [1]. Bahkan oleh konsumen luar negeri, menilai iPhone sebagai produk *overpriced*. Berdasarkan laporan dari Blibli Friends, perbandingan harga antara iPhone dan Android cukup mencolok. Di Amerika Serikat, harga *smartphone* Android berada di kisaran US$100 atau sekitar Rp1,4 juta rupiah, sedangkan iPhone terbaru dijual mulai dari US$999 atau sekitar Rp14 juta. Selain itu, merek *Apple* sering diasosiasikan dengan status sosial dan gaya hidup modern, sehingga meningkatkan daya tariknya di kalangan masyarakat.

Selain terkenal karena spesifikasi dan performanya yang unggul, produk Apple juga sering diasosiasikan dengan status sosial dan gaya hidup modern, yang semakin menambah daya tarik di mata konsumen [2]. Sebagai alternatif, banyak masyarakat memilih untuk membeli produk iPhone *second* yang masih dalam kondisi baik. Pilihan ini memungkinkan mereka tetap bisa menikmati teknologi dan fitur canggih dari produk iPhone tanpa harus mengeluarkan biaya sebesar saat membeli unit baru.

Tokopedia merupakan salah satu *marketplace* terpercaya dan terkemuka peringkat kedua di Indonesia dan mencatat total kunjungan global sebanyak 58 juta berdasarkan data dari SimilarWeb pada Februari tahun 2025. Tokopedia menyediakan berbagai macam produk, salah satunya adalah produk elektronik yaitu iPhone *second* [3]. Pembelian secara daring melalui Tokopedia dapat memberikan kemudahan bagi masyarakat untuk berbelanja tanpa harus mengunjungi toko fisik, sehingga lebih hemat waktu dan efisien. Namun tantangan tetap ada, terutama dalam membangun kepercayaan konsumen kepada penjual terhadap produk *second* seperti iPhone. Kekhawatiran terhadap kualitas dan keaslian barang yang dijual di Tokopedia [4].

Tokopedia menyediakan fitur ulasan yang menjadi salah satu elemen penting dalam membangun kepercayaan pembeli terhadap penjual. Penilaian ini membantu calon pembeli memahami reputasi penjual dan kualitas produk yang ditawarkan konsumen [5]. Namun jumlah ulasan yang begitu banyak di fitur ulasan membuat proses membaca dan menganalisis secara manual menjadi tidak efisien dan memakan banyak waktu. Oleh sebab itu, dibutuhkan metode yang lebih efisien untuk mengelola ulasan dalam jumlah besar, salah satunya melalui penerapan teknik analisis sentimen [6].

Analisis sentimen adalah metode yang digunakan untuk mengolah data teks, seperti ulasan, guna memahami opini atau tanggapan konsumen yang terkandung apakah bersifat positif atau negatif, sehingga dapat dijadikan dasar untuk mengambil keputusan yang lebih tepat terkait suatu prouk, layanan, atau topik tertentu. [7].

Beberapa penelitian sebelumnya telah menyatakan bahwa pendekatan analisis sentimen efektif digunakan untuk menangkap persepsi konsumen terhadap suatu produk. Dalam salah satu penelitian yang membahas ulasan produk masker *Top Brand* di platform Tokopedia dengan menerapkan algoritma *Naïve Bayes*, ditemukan bahwa sebagian besar ulasan memiliki kecenderungan positif. Hasil penelitian tersebut juga memperlihatkan bahwa algoritma *Naïve Bayes* memiliki kemampuan klasifikasi sentimen yang cukup baik, dengan akurasi mencapai 88%.

Penelitian lainnya yang mengkaji analisis sentimen terhadap produk *skincare* dari merek *Skintific* menggunakan data komentar dari media sosial, yang mencakup tanggapan pengguna dan identitas akun mereka. Penelitian ini menerapkan algoritma *Support Vector Machine* (SVM) dengan kernel linear untuk mengklasifikasikan opini pengguna. Hasil penelitian menunjukkan bahwa metode SVM tersebut cukup efektif dalam membedakan sentimen positif dan negatif pada ulasan berbahasa Indonesia, dengan tingkat akurasi mencapai 94%.

Penelitian ini berfokus pada penerapan dua metode klasifikasi, yaitu *Naïve Bayes* dan *Support Vector Machine* (SVM), dalam menganalisis sentimen konsumen terhadap ulasan iPhone *second* yang dijual di Tokopedia. Tujuan dari studi ini adalah untuk menilai dan membandingkan tingkat akurasi kedua algoritma tersebut dalam mengidentifikasi opini pengguna, baik yang bersifat positif maupun negatif. Melalui perbandingan ini, diharapkan dapat diperoleh gambaran yang lebih menyeluruh mengenai efektivitas masing-masing metode dalam memahami preferensi serta kekhawatiran pelanggan terhadap produk iPhone *second* di platform *e-commerce.*

Meski sejumlah penelitian telah membahas analisis sentimen terhadap produk di marketplace menggunakan pendekatan klasifikasi teks berbasis TF-IDF, Naïve Bayes, maupun Support Vector Machine [8],[9], [10], kajian-kajian tersebut umumnya berfokus pada kategori produk yang bersifat umum dan tidak secara spesifik menyoroti sentimen konsumen terhadap perangkat elektronik bekas. Selain itu, penelitian mengenai sentimen pelanggan di Indonesia sebagian besar hanya memanfaatkan satu algoritma atau tanpa membandingkan performa antar model [11]. Hingga saat ini, belum ditemukan penelitian yang secara komprehensif menganalisis sentimen konsumen terhadap iPhone bekas di Tokopedia dengan perbandingan dua algoritma yang paling banyak digunakan dalam text mining, yaitu Naïve Bayes dan Support Vector Machine, serta didukung teknik penyeimbangan data seperti SMOTE yang direkomendasikan untuk dataset tidak seimbang [12]. Hal ini menunjukkan adanya celah penelitian yang penting untuk diisi, terutama pada konteks marketplace Indonesia dan domain produk bekas yang memiliki dinamika opini yang berbeda.

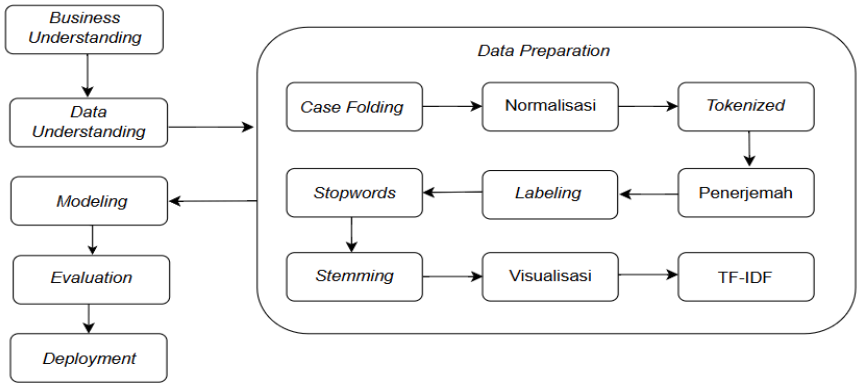
Kebaruan penelitian ini terletak pada integrasi pipeline CRISP-DM yang dipadukan dengan teknik representasi teks TF-IDF [8], penyeimbangan kelas menggunakan SMOTE [12], serta proses penerjemahan teks untuk menormalkan ulasan berbahasa Indonesia sebelum masuk ke tahap klasifikasi. Selain itu, penelitian ini melakukan komparasi performa antara Naïve Bayes—yang dikenal efektif untuk dokumen berfitur tinggi [13] dan Support Vector Machine, salah satu metode paling andal untuk klasifikasi teks dengan fitur besar [9]. Integrasi keseluruhan metode tersebut dalam konteks analisis sentimen produk iPhone bekas di Tokopedia

P-ISSN 2085-4315
E-ISSN 2502-8332

belum pernah dilakukan pada penelitian sebelumnya, sehingga memberikan kontribusi ilmiah berupa pendekatan yang lebih robust untuk menangani karakteristik data ulasan konsumen pada marketplace Indonesia. Sebagai tambahan, penelitian ini juga menerapkan model terbaik ke dalam sebuah sistem prediksi berbasis web sebagai bentuk validasi implementatif, yang tidak ditemukan pada studi-studi terdahulu di domain serupa.

2. METODE PENELITIAN

Penelitian ini dilakukan menggunakan teknik CRISP-DM (Cross-Industry Standard Process for Data Mining) sebagai kerangka kerja dalam pelaksanaan penelitian ini. Gambar 1 merupakan kerangka penelitian yang menjelaskan langkah-langkah dalam penelitian ini.



Gambar 1. Kerangka Penelitian

2.1 *Machine Learning*

Machine Learning merupakan cabang dari kecerdasan buatan yang berfokus pada pengembangan algoritma dan model yang memungkinkan komputer untuk belajar dan membuat prediksi atau keputusan tanpa perlu diprogram secara eksplisit [14].

2.2 Analisis Sentimen

Analisis sentimen merupakan teknik yang digunakan untuk memahami opini, perasaan, atau penilaian yang diberikan oleh pengguna internet melalui media sosial [15]. Hal ini sangat relevan dalam dunia bisnis, di mana analisis sentimen diterapkan untuk menganalisis ulasan produk, layanan pelanggan, dan tanggapan konsumen terhadap suatu produk atau merek.

2.3 *Web Scraping*

Web Scraping merupakan metode pengumpulan data secara otomatis dari situs *web* menggunakan teknik-teknik tertentu untuk mengekstrak informasi yang tersedia di halaman *web* [16]. Teknik *Web Scraping* dapat memanfaatkan bahasa pemrograman seperti Python, yang sangat populer karena memiliki pustaka-pustaka seperti *BeautifulSoup*, *Scrapy,* dan *Selenium*.

2.4 *Naïve Bayes*

Naïve Bayes merupakan algoritma pembelajaran mesin yang mengandalkan *Teorema Bayes* untuk memprediksi kategori dari suatu data. Algoritma ini beroperasi dengan asumsi bahwa setiap fitur dalam data saling independen, yang membuat perhitungan probabilitas menjadi lebih efisien dan lebih sederhana. Berikut perhitungan menggunakan *Teorema Bayes* dan distribusi *multinomial*, yang dijelaskan melalui rumus berikut ini [17].

$$P(C_k|x) = \frac{P(C_k) \prod_{i=1}^{n} P(x_1|C_k)}{P(x)} \quad (1)$$

Keterangan:

$P(C_k|x)$ = Probabilitas kelas C_k C_k diberikan fitur *X* (posterior).

P-ISSN 2085-4315
E-ISSN 2502-8332

$P(C_k)$ = Kelas tertentu yang ingin diidentifikasi dalam data X.

$P(x_1|C_k)$ = Kelas yang dimaksud untuk dianalisis.

$P(x)$ = Probabilitas total dari semua kemungkinan fitur

2.5 *Support Vector Machine* (SVM)

Support Vector Machine (SVM) merupakan algoritma pembelajaran mesin yang beroperasi dengan cara mencari *hyperplane* terbaik yang dapat memisahkan dua kelas data dengan margin terbesar di dalam ruang fitur [18]. Dalam SVM, titik data yang paling krusial adalah *support vectors,* yaitu titik-titik yang paling dekat dengan *hyperplane* pemisah [19]. *Support Vectore Machine, hyperplane* pemisah dijelaskan melalui persamaan berikut:

$$w_0 + w_1 x_1 + w_2 x_2 > 0 \quad (2)$$

Oleh karena itu, setiap titik yang terletak di atas hyperplane pemisah memenuhi kondisi:

$$w_0 + w_1 x_1 + w_2 x_2 < 0 \quad (3)$$

Sedangkan setiap titik yang terletak di bawah hyperplane pemisah memenuhi kondisi:

$$w_0 + w_1 x_1 + w_2 x_2 > 0 \quad (4)$$

Pada tahap pra-pemrosesan data, seluruh ulasan berbahasa Indonesia diterjemahkan terlebih dahulu ke dalam bahasa Inggris sebelum dilakukan analisis sentimen menggunakan VADER. Langkah ini diperlukan karena VADER merupakan rule-based sentiment analyzer yang secara eksplisit dikembangkan untuk teks berbahasa Inggris, dengan kamus leksikal dan aturan sintaksis yang didesain berdasarkan ekspresi linguistik penutur asli bahasa Inggris. Hutto dan Gilbert menegaskan bahwa VADER mengandalkan struktur bahasa Inggris dalam penanganan intensifier, negation, idiom, slang, hingga penekanan berbasis tanda baca dan kapitalisasi, sehingga performanya optimal hanya ketika diterapkan pada teks berbahasa Inggris [20]. Oleh karena itu, penggunaan VADER secara langsung pada teks Indonesia berisiko menghasilkan bias karena keterbatasan cakupan leksikal dan ketidaksesuaian aturan sintaksis. Dengan menerjemahkan ulasan ke bahasa Inggris, model dapat mengenali polaritas sentimen secara lebih akurat sesuai karakteristik lexicon VADER.

Dataset ulasan iPhone bekas pada penelitian ini menunjukkan ketidakseimbangan kelas yang cukup signifikan, di mana proporsi ulasan positif jauh lebih dominan dibandingkan ulasan negatif. Kondisi ini berpotensi menyebabkan model klasifikasi belajar secara bias dan cenderung mengabaikan pola dari kelas minoritas, sehingga menghasilkan metrik evaluasi yang tidak representatif. Untuk mengatasi hal tersebut, penelitian ini menerapkan SMOTE (Synthetic Minority Over-sampling Technique), yaitu teknik oversampling berbasis interpolasi yang menghasilkan sampel sintetis baru untuk kelas minoritas. Menurut Chawla et al., SMOTE terbukti efektif dalam meningkatkan kemampuan model untuk mempelajari distribusi kelas minoritas tanpa menimbulkan overfitting sebagaimana terjadi pada metode oversampling berbasis duplikasi [12]. Penelitian lanjutan oleh Fernández et al. juga menunjukkan bahwa SMOTE merupakan pendekatan yang stabil dan unggul dalam menangani ketidakseimbangan data pada berbagai tugas klasifikasi, termasuk pemrosesan teks [21]. Dengan demikian, penggunaan SMOTE dalam penelitian ini memberikan dasar metodologis yang kuat untuk meningkatkan kinerja dan keadilan model (fair learning) dalam proses klasifikasi sentimen.

3. HASIL DAN PEMBAHASAN

3.1. *Business Understanding*

Tren pembelian iPhone bekas secara online menunjukkan peningkatan, khususnya melalui platform Tokopedia. Dari hasil survei awal, ditemukan bahwa 58,1% responden memilih Tokopedia karena dinilai memiliki tingkat kepercayaan yang tinggi. Fakta ini memperlihatkan bahwa rasa percaya terhadap platform sangat memengaruhi keputusan konsumen dalam membeli perangkat elektronik bekas secara daring. Penelitian ini berupaya mengevaluasi persepsi konsumen terhadap produk dan layanan melalui analisis sentimen pada ulasan pengguna. Proses klasifikasi dilakukan dengan pendekatan machine learning menggunakan algoritma Naïve Bayes dan Support Vector Machine (SVM) yang dikenal memiliki performa baik dalam pengolahan teks.

P-ISSN 2085-4315
E-ISSN 2502-8332

Temuan dari penelitian ini diharapkan dapat memberikan wawasan kepada pelaku bisnis dan platform e-commerce dalam memahami pola sentimen konsumen, merancang strategi pemasaran yang lebih terarah, serta memberikan dukungan pengambilan keputusan bagi calon pembeli berdasarkan kecenderungan ulasan yang tersedia.

3.2 *Data Understanding*

Data ulasan pelanggan diperoleh dari empat toko yang menjual iPhone second di Tokopedia, yakni Boboy Store, JoyTaroGadgetOfficial, Maxsense Store, dan Yunishop00. Pengambilan data dilakukan secara otomatis menggunakan metode web scraping dengan bantuan library Selenium, BeautifulSoup, dan Pandas.

Ulasan yang dianalisis berasal dari pembeli yang memposting ulasan dalam rentang waktu Januari 2023 hingga Januari 2025, dengan total 1.863 baris data dalam format CSV. Gambar 2 menampilkan dataset dari salah satu toko hasil teknik web scraping, yaitu Toko Boboy Store.

	Username	Product	Tanggal	Rating	Ulasan
0	Ranita	Iphone 13 128Gb / 256Gb / 512Gb Bekas Fullset	1 hari lalu	5	toko sangat bagus udah biasa langganan disini
1	Alexander	Iphone 14 Pro 128Gb / 256Gb / 512Gb / 1TB Seco...	1 hari lalu	5	barang bagus, imei tembus, hape mulus. semoga ...
2	Erdy	Iphone 12 Pro 128gb / 256gb /512gb Bekas Fullset	1 hari lalu	5	Sangat memuaskan barang dari bosku yg 1 ini ,,...
3	D***g	Iphone 13 Mini 128Gb / 256Gb / 512Gb Second Im...	2 hari lalu	5	Makasih banyak bosku, sudah sampai barangnya a...
4	upi	Iphone 13 128Gb / 256Gb / 512Gb Bekas Fullset	2 hari lalu	5	Sudah yg ketiga kalinya beli disini, hasilnya ...

Gambar 2. Hasil Scraping Ulasan pada Toko Boboystore

Proses penyaringan dilakukan dengan Python untuk memastikan hanya ulasan produk iPhone second yang digunakan dalam analisis. Hasil penyaringan ditampilkan pada Gambar 3.

	Username	Product	Tanggal	Rating	Ulasan
13	MUSTARING	Apple Charger Original 20 W Usb C Power Adapto...	3 minggu lalu	5	Barang sangat baik
17	Andhika	Ipad Air 2 16GB 32GB 64GB 128GB Wifi Only Wifi...	3 minggu lalu	5	Admin ramah, toko yang terjamin
18	dicky	iPad 8 7 6 5 128GB 32GB WiFi Only Wifi Cell Fu...	3 minggu lalu	5	alus madep pokona mah
32	Bagas	iPad Pro M1 2021 11 & 12.9 inch WiFi Cellular ...	1 bulan lalu	5	Mantap banget sihh masih aman sampe skarang so...
39	Adi	SEIN Samsung Galaxy S10 Plus 12/1TB 8/512GB 8/...	1 bulan lalu	5	recomended seller

Gambar 3. Hasil Produk Bukan iPhone Second

Data dari keempat toko yang telah disaring digabungkan menjadi satu dataset menggunakan Python, kemudian disimpan dalam format CSV. Gambar 4 menampilkan hasil akhir proses tersebut.

	Username	Product	Tanggal	Rating	Ulasan
0	N***a	iphone 11 64gb 128gb second fullset mulus terawat	Hari ini	5	BH suntikan, baterai sudah pernah diganti... Tid...
1	Irfan	iPhone 11 Second IMEI Terdaftar BC (BeaCukai)...	1 hari lalu	5	Barang sesuai pesanan, mulus banget, IMEI terd...
2	Muhammad	iPhone XR 64gb 128gb second fullset ex inter	2 hari lalu	5	baterai cepat panas. boros pemakaian, sehari s...
3	A***t	iphone 13 pro max 256gb second fullset mulus e...	4 hari lalu	5	Seller Fast Respon, Produk sesuai deskripsi, o...
4	W***y	iPhone 15 pro 256gb second fullset mulus terawat	5 hari lalu	5	Kamera jernih bgt aku pake sampai sekarang ser...
...	...	...	...	...	...
1858	D***i	IPHONE 11 GREEN 128GB | MULUS LIKE NEW | 100% ...	Lebih dari 1 tahun lalu	5	pengiriman cepat sampai, toko bertanggung jawa...
1859	A***a	IPHONE 11 PRO MAX SILVER 256GB 100% ORIGINAL |...	Lebih dari 1 tahun lalu	5	Barangnya original ,, Sinyal tidak terblokir ...
1860	Yassina	IPHONE 12 PRO MAX PACIFIC BLUE 256GB MULUS LIK...	Lebih dari 1 tahun lalu	5	Barang mulus, bh 97, pengirimannya cepat , pac...
1861	eka	IPHONE 12 PRO MAX PACIFIC BLUE 512GB ORIGINAL ...	Lebih dari 1 tahun lalu	5	barang ok, msh bagus dan mulus kaya baru, toko...
1862	Hermawan	IPHONE 13 PRO SIERRA BLUE 128GB MULUS LIKE NEW...	Lebih dari 1 tahun lalu	5	Barangnya original , Sinyal aman tidak berma...

1863 rows × 5 columns

Gambar 4. Hasil Penggabungan Dataset

3.3 *Data Preparation*

Tahap ini meliputi proses awal dalam menyiapkan data ulasan iPhone second di Tokopedia sebelum memasuki tahap pemodelan.

1. Case Folding

Pada tahap case folding, data dibersihkan untuk meningkatkan kualitasnya. Langkah ini mencakup

penghapusan emoji, angka, spasi ganda, karakter tidak relevan, serta spasi di sekitar tanda baca. Nilai null dan kolom yang tidak relevan dihapus, sementara kolom Rating dan Ulasan dipertahankan karena dianggap penting untuk analisis sentimen.

2. Tokenized

Pada tahap tokenisasi, teks dipecah menjadi unit-unit kecil seperti kata atau frasa untuk memudahkan pemrosesan lanjutan. Perbandingan sebelum dan sesudah tokenisasi ditampilkan pada Gambar 5.

Ulasan	Tokenized
bh disuntik tidak ril btrai cpt panas ceritany...	[bh, disuntik, tidak, ril, btrai, cpt, panas, ...
sesuai pesanan mulus bget imei terdaftar sinya...	[sesuai, pesanan, mulus, bget, imei, terdaftar...
baterai cept panas bagian microphone nya berke...	[baterai, cept, panas, bagian, microphone, nya...
seller fast respon produk sesuai deskripsi ove...	[seller, fast, respon, produk, sesuai, deskrip...
kamera jernih bgt aku pake sampai sekarang ser...	[kamera, jernih, bgt, aku, pake, sampai, sekar...

Gambar 5. Perbedaan Teks antara Kolom Ulasan dan Kolom Tokenized

3. Normalisasi

Tahap normalisasi dilakukan setelah tokenisasi, dengan tujuan menyempurnakan data melalui koreksi kata tidak baku, singkatan, dan kesalahan penulisan. Tabel 1 menampilkan perbandingan kata sebelum dan sesudah normalisasi.

Tabel 1. Perbandingan Kata Sebelum dan Setelah Normalisasi

Sebelum	Sesudah	Sebelum	Sesudah
amboxing	unboxing	belanj	belanja
batre	baterai	cc	coba-coba
bh	battery health	cpt	cepat

4. Penerjamahan

Tahap penerjemahan dilakukan untuk mempermudah proses pelabelan sentimen dengan mengonversi teks ulasan ke dalam bahasa Inggris. Langkah ini juga memungkinkan pemanfaatan alat analisis yang lebih optimal karena banyak di antaranya dirancang khusus untuk bahasa Inggris. Gambar 6 menunjukkan perbandingan antara kolom Ulasan asli dan kolom English_Review hasil terjemahan.

Ulasan	English_Review
baterai health suntikkan tidak asli baterai ce...	the health battery injection is not authentic,...
sesuai pesanan mulus banget imei terdaftar sin...	according to the order is very smooth imei is ...
baterai cepat panas bagian microphone nya berk...	the battery is hot, the microphone is constrai...
seller cepat respon produk sesuai deskripsi ke...	seller fast product response according to the ...

Gambar 6. Hasil Penerjemahan ke Kolom English_Review

5. Labelling

Pada tahap pelabelan, ulasan berbahasa Inggris diklasifikasikan ke dalam dua kategori sentimen: positif dan negatif. Proses ini menggunakan pustaka VADER (Valence Aware Dictionary and Sentiment Reasoner), yang menentukan polaritas berdasarkan skor komposit. Ulasan dengan skor ≥ 0 dilabeli sebagai positif, sedangkan skor < 0 dianggap negatif. Untuk meningkatkan akurasi analisis sesuai konteks, peneliti menambahkan kosakata khusus (custom lexicon) ke dalam VADER. Daftar kata negatif tambahan ditampilkan pada Tabel 2.

Tabel 2 Daftar Kata-Kata dengan Sentiment Negatif

Kata	Skor
blisters	-3.0
hot_battery	-3.0
injection	-3.0

Selanjutnya dilakukan pengecekan manual untuk memastikan kesesuaian hasil pelabelan, karena tidak semua ulasan iPhone second di Tokopedia terklasifikasi dengan tepat secara otomatis. Gambar 7 menunjukkan hasil

P-ISSN 2085-4315
E-ISSN 2502-8332

akhir pelabelan berdasarkan polaritas sentimen masing-masing ulasan.

Rating	Ulasan	English_Review	Label
1	baterai health suntikkan tidak asli baterai ce...	the health battery injection is not authentic,...	negatif
5	sesuai pesanan mulus banget imei terdaftar sin...	according to the order is very smooth imei is ...	positif
2	baterai cepat panas bagian microphone nya berk...	the battery is hot, the microphone is constrai...	negatif
5	seller cepat respon produk sesuai deskripsi ke...	seller fast product response according to the ...	positif
5	kamera jernih banget aku pakai sampai sekarang...	i really use clear camera until now i often us...	positif

Gambar 7. Hasil dari Proses Pelabelan

6. Stopwords

Kata-kata yang tidak memiliki nilai informasi penting dihapus menggunakan proses stopwords removal. Hasil implementasinya ditampilkan pada Tabel 3.

Tabel 3. Hasil dari Proses Pelabelan

Sebelum *Stopwords*	Sesudah *Stopwords*
minus nya hanya sinyal bulanan imei tidak terdaftar bea cukai tapi semoga aja bertahan lama *thanks*	*minus* hanya sinyal bulanan imei tidak terdaftar bea cukai semoga bertahan lama.

7. Stemming

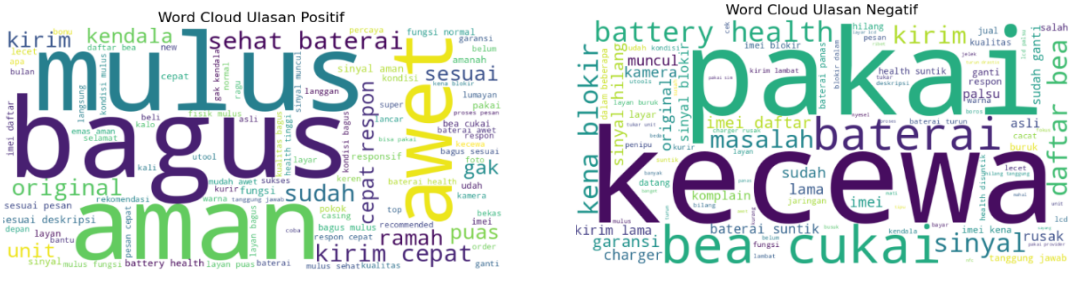
Proses stemming dilakukan untuk mengubah setiap kata pada kolom ulasan ke bentuk dasarnya. Hasil dari proses ini ditampilkan pada Tabel 9.

Tabel 4. Hasil dari Proses Stemming

Sebelum *Stemming*	Sesudah *Stemming*
minus nya hanya sinyal bulanan imei tidak terdaftar bea cukai tapi semoga aja bertahan lama thanks seller.	minus hanya sinyal bulan imei tidak daftar bea cukai semoga tahan lama.

8. Visualisasi

Tahap selanjutnya adalah visualisasi data menggunakan *Word Cloud* untuk menggambarkan kata-kata yang paling sering muncul dalam ulasan iPhone *second* di Tokopedia, berdasarkan klasifikasi sentimen positif dan negatif. Visualisasi ini bertujuan untuk mengidentifikasi pola kepuasan dan ketidakpuasan konsumen, sehingga dapat memberikan gambaran umum bagi calon pembeli mengenai aspek-aspek yang sering menjadi perhatian. Ukuran kata yang lebih besar menunjukkan frekuensi kemunculan yang lebih tinggi dalam dataset. Gambar 8 menampilkan perbandingan *Word Cloud* antara sentimen positif dan negatif.



(a) (b)

Gambar 8. *Word Cloud* dengan (a) Sentimen Positif dan (b) Sentimen Negatif

9. Term Frequency (TF) dan Inverse Document Frequency (IDF)

Sebelum dilakukan transformasi menggunakan TF-IDF, data ulasan iPhone second dibagi menjadi tiga subset: data latih (70%), serta data validasi dan data uji (masing-masing dari sisa 30%). Pembagian ini bertujuan untuk mengevaluasi performa model secara menyeluruh. Proses TF-IDF dilakukan pada data latih untuk membangun pemahaman awal model terhadap kata-kata dan bobot pentingnya dalam teks. Tabel 5 menampilkan hasil perhitungan TF-IDF pada dua contoh ulasan.

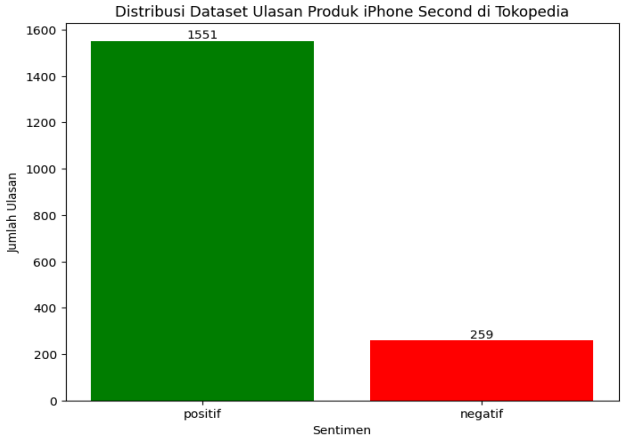
JURNAL ILMIAH FIFO
Vol. 17 No. 2 (2025) 212-224
DOI: 10.22441/fifo.2025.v17i2.010
P-ISSN 2085-4315
E-ISSN 2502-8332

Tabel 5. Hasil perhitungan TF-IDF

Kata	TF		TF Normalisasi		DF	IDF	TF-IDF	
	U1	U2	U1	U2			U1	U2
sesuai	1	0	1/6	0	1	log(2/1) = 0,301	0,301 × (1/6) = 0,050	0
pesan	1	0	1/6	0	1	0,301	0,050	0
mulus	1	0	1/6	0	1	0,301	0,050	0
imei	1	1	1/6	1/7	2	log(2/2) = 0	0	0
daftar	1	0	1/6	0	1	0,301	0,050	0

3.4 Modeling

Sebelum membangun model Machine Learning, perlu diiawali dengan analisis distribusi sentimen pada dataset ulasan iPhone second di Tokopedia yang telah melalui proses TF-IDF.



Gambar 9. Distribusi Sentimen Data Ulasan iPhone Second di Tokopedia

Gambar 9 menunjukkan distribusi kelas yang tidak seimbang, di mana ulasan negatif jumlahnya lebih sedikit, sehingga dapat memengaruhi kemampuan model dalam mengenali kelas tersebut. Untuk mengatasi hal ini, digunakan teknik oversampling dengan metode SMOTE. Selanjutnya, model klasifikasi sentimen dibangun menggunakan algoritma Naïve Bayes dan Support Vector Machine (SVM). Kedua algoritma dilatih pada data yang sama dan dibandingkan performanya untuk menentukan model terbaik.

1. Naïve Bayes

Model klasifikasi sentimen menggunakan algoritma Multinomial Naïve Bayes berhasil mencapai akurasi sebesar 93% dalam mengelompokkan ulasan iPhone second di Tokopedia ke dalam kelas positif dan negatif. Hasil ini menunjukkan bahwa algoritma tersebut mampu mengenali pola sentimen dengan cukup baik, meskipun masih memiliki keterbatasan dalam menangani konteks yang lebih kompleks pada data teks.

2. Support Vector Machine (SVM)

Model klasifikasi sentimen menggunakan algoritma Support Vector Machine (SVM) dengan threshold 0.2 menghasilkan akurasi sebesar 96%. Angka ini menunjukkan bahwa SVM mampu mengidentifikasi sentimen ulasan iPhone second di Tokopedia dengan tingkat ketepatan yang tinggi. Kinerja yang unggul ini mencerminkan kemampuan SVM dalam memisahkan kelas positif dan negatif secara lebih optimal dibandingkan model Multinomial Naïve Bayes, terutama dalam menangani dimensi fitur yang tinggi dari representasi teks.

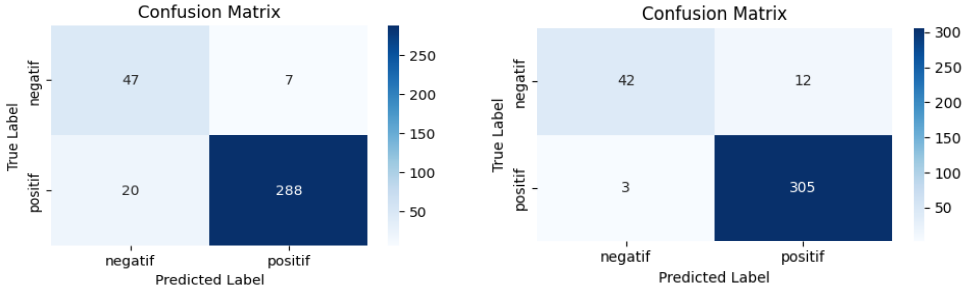
3.5 Evaluation

Berdasarkan hasil evaluasi, algoritma Naïve Bayes dan Support Vector Machine (SVM) sama-sama menunjukkan performa yang cukup baik dalam klasifikasi sentimen ulasan iPhone second di Tokopedia. Meskipun demikian, masing-masing memiliki karakteristik dan keunggulan tersendiri yang memengaruhi pemilihan model sesuai dengan kebutuhan analisis.

JURNAL ILMIAH FIFO
Vol. 17 No. 2 (2025) 212-224
DOI: 10.22441/fifo.2025.v17i2.010

P-ISSN 2085-4315
E-ISSN 2502-8332

Gambar 10 menunjukkan confusion matrix dari kedua model, yang merepresentasikan akurasi klasifikasi ulasan ke dalam sentimen positif dan negatif.



(a) (b)

Gambar 10 *Confusion Matrix* dengan model (a) *Naïve Bayes* dan (b) *Support Vector Machine* (SVM)

Model Naïve Bayes menghasilkan 27 kesalahan klasifikasi, terdiri dari 19 ulasan positif yang keliru diklasifikasikan sebagai negatif dan 8 ulasan negatif yang diprediksi positif. Sementara itu, model SVM hanya mencatat 16 kesalahan, dengan 2 pada ulasan positif dan 14 pada ulasan negatif. Hasil ini menunjukkan bahwa SVM memiliki tingkat akurasi yang lebih tinggi dan lebih andal, khususnya dalam mengenali sentimen positif. Dengan jumlah kesalahan yang lebih rendah, SVM terbukti lebih stabil dalam menghadapi ketidakseimbangan data. Ringkasan evaluasi dari kedua algoritma disajikan pada Tabel 6.

Tabel 6. Perbandingan evaluasi kedua model yaitu Naïve Bayes dan SVM.

Algoritma	Accuracy	Precision	Recall	F1-Score
Naïve Bayes	0.93	0.84	0.90	0.87
Support Vector Machine	0.96	0.95	0.88	0.91

Support Vector Machine (SVM) dipilih sebagai algoritma terbaik karena menunjukkan performa yang lebih unggul dalam mengklasifikasikan sentimen positif dan negatif pada ulasan iPhone second di Tokopedia. SVM mencapai akurasi sebesar 95%, lebih tinggi dibandingkan Naïve Bayes yang mencatatkan akurasi 92%. Keunggulan ini semakin terlihat pada hasil pengujian model, di mana SVM hanya menghasilkan 9 kesalahan dari total 362 ulasan, sehingga mencapai akurasi 97% (Gambar 11a). Sebaliknya, Naïve Bayes melakukan 37 kesalahan klasifikasi pada data yang sama, dengan akurasi sebesar 90% (Gambar 11b). Temuan ini menunjukkan bahwa model Support Vector Machine lebih konsisten dan akurat dalam memproses data ulasan yang tidak seimbang.

```
Evaluasi Model Support Vector Machine dengan Data Uji
Akurasi     : 0.9751
Precision   : 0.9748
Recall      : 0.9751
F1-Score    : 0.9748
Prediksi Salah: 9 data
```

(a)

```
Evaluasi Model Naive Bayes dengan Data Uji
Akurasi     : 0.8978
Precision   : 0.9285
Recall      : 0.8978
F1-Score    : 0.9061
Prediksi Salah: 37 data
```

(b)

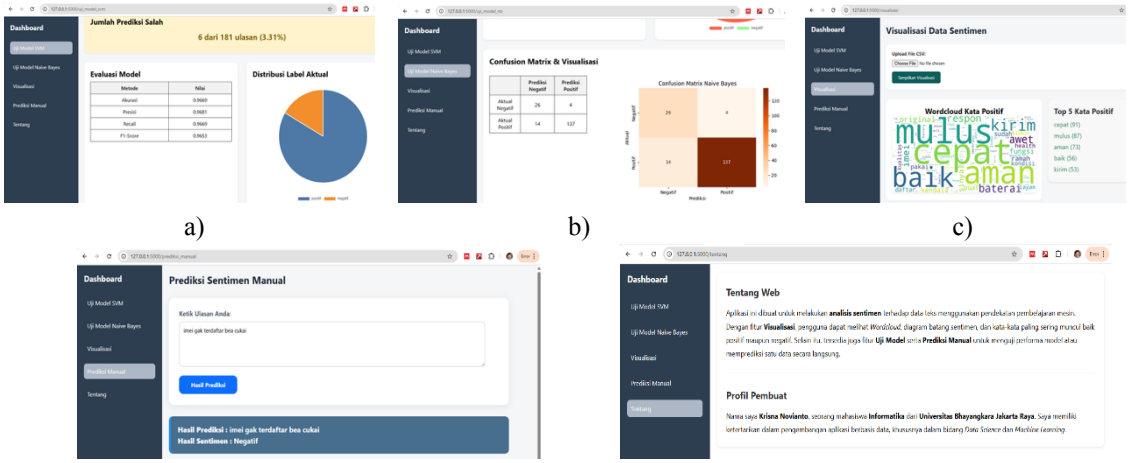
Gambar 11. Hasil Evaluasi (a) Algoritma SVM dengan Data Validasi dan (b) Evaluasi Algoritma Naïve Bayes dengan Data Validasi

3.6 Deployment

Pada tahap *deployment*, model *Support Vector Machine* diimplementasikan di dalam aplikasi *web* bernama analisis sentimen iPhone *second* di Tokopedia menggunakan *framework Flask*. File TF-IDF yang dihasilkan dari pemrosesan data latih ulasan produk iPhone *second* juga digunakan dalam sistem ini. *Web* ini menyediakan empat fitur yaitu: uji model SVM, uji model *Naïve Bayes*, visualisasi, prediksi manual, dan tentang *web.* Gambar 13 merupakan seluruh tampilan fitur dari *web*.

JURNAL ILMIAH FIFO
Vol. 17 No. 2 (2025) 212-224
DOI: 10.22441/fifo.2025.v17i2.010

P-ISSN 2085-4315
E-ISSN 2502-8332

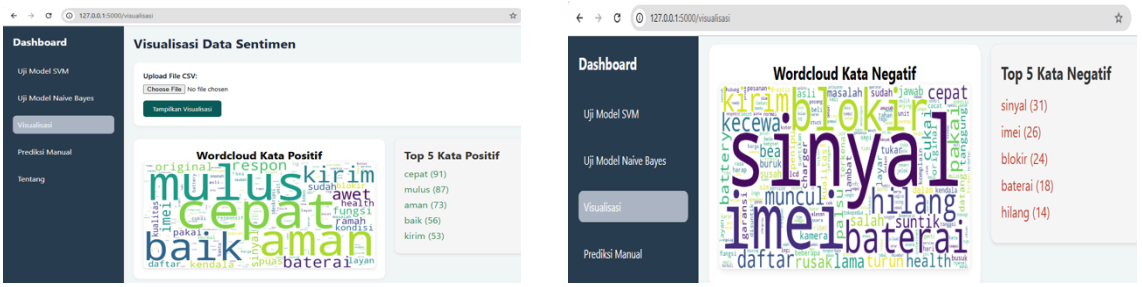


a) b) c)

d) e)

Gambar 12. Tampilan *web* sentimen analisis produk iPhone *second* di Tokopedia pada (a) fitur uji model SVM, (b) Fitur uji model *Naïve Bayes*, (c) Fitur Visualisasi (d) Fitur prediksi manual, (d) Fitur tentang

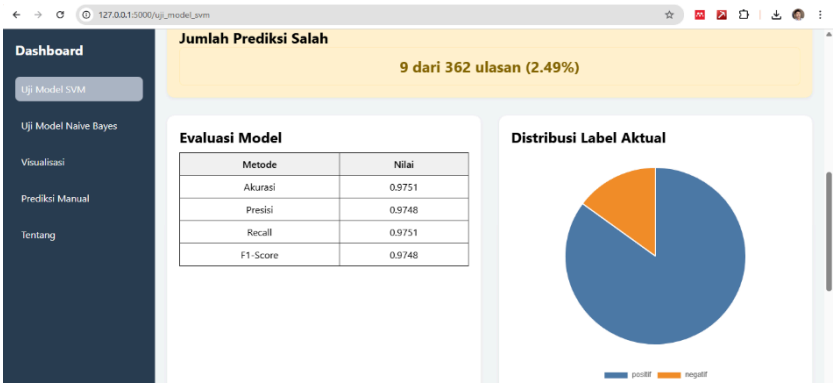
Pada fitur prediksi manual, pengguna mengetik langsung kalimat ulasan yang berkaitan dengan produk iPhone *second* di Tokopedia. Kalimat ulasan yang diketik pengguna adalah: *"imei gak terdaftar di bea cukai"*, menghasilkan bahwa kalimat tersebut memiliki sentimen "negatif", dapat dilihat pada Gambar 13 (b). Sementara itu, dilakukan pengunggah data latih ke fitur visualisasi sebagaimana pada Gambar 14 (a) dan (b), menghasilkan tampilan *word cloud* dan lima daftar kata yang sering muncul pada masing-masing sentimen.



a) b)

Gambar 13. Hasil Visualisasi dari data latih (a) sentimen positif (b) sentimen Negatif

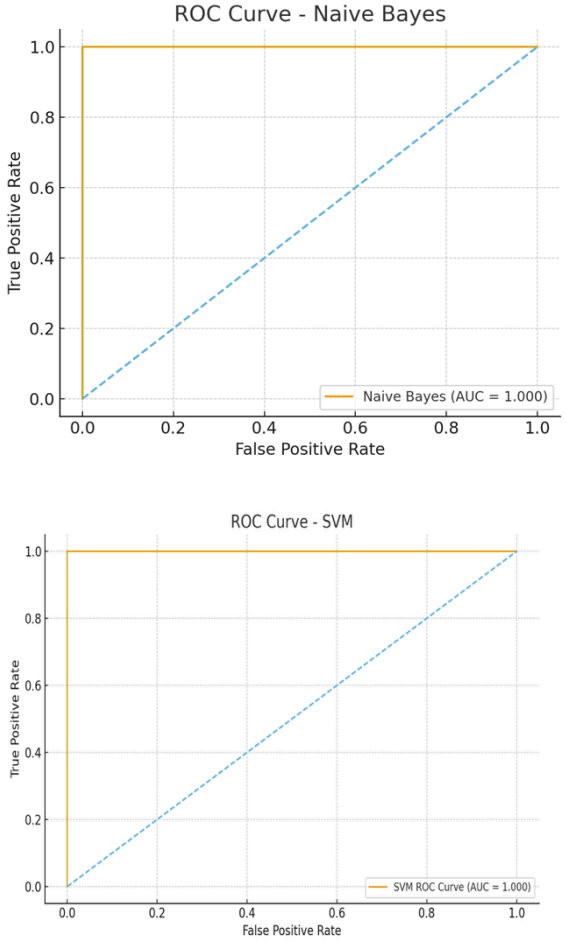
Pada fitur uji model SVM, dilakukan pengujian terhadap 362 ulasan. Hasilnya, model salah memprediksi 9 ulasan dan mencapai akurasi sebesar 97%. Capaian ini menunjukkan bahwa *Support Vector Machine* merupakan model dengan performa terbaik. Rincian hasil ditampilkan pada Gambar 15.



Gambar 14. Hasil Prediksi Model SVM terhadap data latih

JURNAL ILMIAH FIFO
Vol. 17 No. 2 (2025) 212-224
DOI: 10.22441/fifo.2025.v17i2.010
P-ISSN 2085-4315
E-ISSN 2502-8332

Analisis performa model juga dilakukan melalui evaluasi Receiver Operating Characteristic (ROC) untuk membandingkan kemampuan diskriminatif antara Naïve Bayes dan Support Vector Machine. Berdasarkan kurva ROC yang dihasilkan, model SVM menunjukkan Area Under Curve (AUC) yang lebih tinggi dibandingkan Naïve Bayes, yang mengindikasikan bahwa SVM memiliki kemampuan pemisahan kelas yang lebih baik pada semua threshold. Sementara Naïve Bayes cenderung stabil namun kurang optimal pada data ulasan yang memiliki distribusi kata tidak seimbang, SVM mampu mempertahankan sensitivitas dan spesifisitas yang lebih tinggi, terutama pada kelas negatif yang jumlahnya lebih sedikit (Gambar 15).



Gambar 5. Grafik ROC Naive Bayes dan SVM

Hasil analisis kesalahan menunjukkan bahwa sebagian ulasan yang secara manual dikategorikan sebagai neutral cenderung diprediksi sebagai negatif oleh kedua model, terutama Naïve Bayes. Fenomena ini muncul karena dua faktor utama. Pertama, ulasan neutral dalam konteks pembelian produk bekas sering mengandung kata yang menggambarkan kondisi apa adanya ("lumayan", "sesuai harga", "normal"), yang oleh model dibaca sebagai sinyal ketidakpuasan ringan. Kedua, ketidakseimbangan kelas menyebabkan model lebih sensitif terhadap kata-kata yang bernada negatif meskipun konteksnya netral. Pada kasus lain, ulasan neutral yang hanya berupa deskripsi faktual ("barang sudah diterima") juga berpotensi salah diklasifikasikan karena tidak memiliki kata bernada positif sehingga model mengarahkannya ke kelas lain berdasarkan probabilitas dominan.

Temuan penelitian ini sejalan dengan studi-studi sebelumnya yang menunjukkan bahwa SVM cenderung memberikan performa terbaik dalam klasifikasi sentimen berbasis TF-IDF. Penelitian oleh Jianqiang et al. (2018) melaporkan bahwa SVM unggul dibandingkan Naïve Bayes dan KNN pada data komentar publik, khususnya pada fitur TF-IDF berdimensi tinggi [22]. Penelitian lain oleh Ghifari et al. (2025) juga menunjukkan bahwa SVM mampu memberikan akurasi lebih tinggi Studi ini membandingkan dua metode pemrosesan teks, yaitu Convolutional Recurrent Neural Network (CRNN) sebagai model deep learning dan Support Vector Machine (SVM) sebagai model machine learning klasik. Penelitian ini menilai kinerja kedua model dalam

P-ISSN 2085-4315
E-ISSN 2502-8332

mengklasifikasi opini positif, negatif, atau netral, sekaligus menunjukkan bagaimana kombinasi fitur tekstual dan arsitektur model mempengaruhi akurasi. Hasilnya memberikan gambaran mengenai persepsi publik terhadap performa pelatih nasional serta menunjukkan model mana yang lebih efektif untuk tugas analisis sentimen pada konteks sepak bola Indonesia. [23]. Selain itu, studi oleh Almeida et al. (2011) menemukan bahwa model probabilistik seperti Naïve Bayes cenderung sensitif terhadap ketidakseimbangan kelas dan noise, sedangkan SVM lebih konsisten pada berbagai dataset teks [24]. Dengan demikian, kinerja SVM yang lebih baik pada ulasan iPhone bekas di Tokopedia konsisten dengan tren penelitian sebelumnya.

4. KESIMPULAN

Penelitian ini menunjukkan bahwa calon pembeli iPhone second di Tokopedia berpotensi menghadapi berbagai risiko, seperti kondisi produk yang tidak sesuai deskripsi, kerusakan pada komponen penting, serta isu legalitas seperti IMEI yang tidak terdaftar. Informasi tersebut umumnya baru terungkap melalui ulasan pengguna. Analisis sentimen terhadap ulasan memberikan gambaran yang bermanfaat bagi konsumen. Ulasan positif cenderung menyoroti kondisi fisik yang baik dan pengiriman cepat, sedangkan ulasan negatif berkaitan dengan masalah teknis dan keaslian perangkat. Fitur word cloud dan prediksi otomatis membantu menyederhanakan informasi tersebut agar lebih mudah dipahami. Dari dua algoritma yang digunakan, Support Vector Machine menunjukkan performa paling unggul dengan akurasi 96%, dibandingkan Naïve Bayes yang mencapai 93%. Hasil ini menjadikan SVM sebagai algoritma yang direkomendasikan untuk klasifikasi sentimen dalam konteks ini.

DAFTAR PUSTAKA

[1] I. Hakim and L. S. Harahap, “Analisis Sentimen Terhadap Iphone 16 Pada Data Twitter Menggunakan Orange Data Mining,” *Jurnal Multidisiplin Saintek*, vol. 5, no. 3, 2024.

[2] K. N’da, J. Ge, S. J. Ren, and J. Wang, “What matters for international consumers ’ choice preferences for smartphones : Evidence from a cross-border ecommerce platform,” pp. 1–20, 2023.

[3] D. Nasikah and S. Fuadi, “Pengaruh Persepsi Keamanan, Kemudahan Bertransaksi, Kepercayaan Konsumen, Kualitas Produk Dan Harga Terhadap Keputusan Pembelian Pada Marketplace Tokopedia,” *Jurnal Manajemen*, vol. 16, no. 1, pp. 162–173, 2022.

[4] N. A. Salsabilaa, U. Sa’adah, and F. Fauzi, “Analisis Sentimen Pada Ulasan Aplikasi Tokopedia Menggunakan Klasifikasi Naïve Bayes,” *PRISMA, Prosiding Seminar Nasional Matematika*, pp. 44–51, 2024.

[5] R. Hidayansyah, “Pengaruh Ulasan Dan Penilaian Konsumen Terhadap Keputusan Pembelian Pada Marketplace Tokopedia Di Indonesia,” *Digital Repository UNILA*, 2024.

[6] R. T. Handayanto, Herlawati, P. D. Atika, F. N. Khasanah, A. Y. P. Yusuf, and D. Y. Septia, “Analisis Sentimen Pada Situs Google Review dengan Naïve Bayes dan Support Vector Machine,” *Jurnal Komtika (Komputasi dan Informatika)*, vol. 5, no. 2, pp. 153–163, 2021.

[7] Herlawati, D. B. Srisulistiowati, S. C. Agustin, P. H. Syafina, N. Rachmatin, S. Setiawati, and *, “Metode Naïve Bayes dan Support Vector Machine untuk Mengolah Sentimen Ulasan dan Komentar di Platform Digital,” *Journal of Students’ Research in Computer Science (JSRCS)*, vol. 5, no. 2, pp. 197–212, 2024.

[8] G. Salton and C. Buckley, “Term-weighting approaches in automatic text retrieval,” *Information Processing & Management*, vol. 24, no. 5, pp. 513–523, 1988.

[9] T. Joachims, “Text categorization with support vector machines: Learning with many relevant features,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 1398, pp. 137–142, 1998.

[10] P. Sasikala and L. Mary Immaculate Sheela, “Sentiment analysis of online product reviews using DLMNN and future prediction of online product using IANFIS,” *Journal of Big Data*, vol. 7, no. 1, 2020.

[11] M. Fikri and R. Sarno, “A comparative study of sentiment analysis using SVM and Senti Word Net,” *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 13, no. 3, pp. 902–909, 2019.

[12] R. Peranginangin, E. J. G. Harianja, I. K. Jaya, and B. Rumahorbo, “Penerapan Algoritma Safe-Level-Smote Untuk Peningkatan Nilai G-Mean Dalam Klasifikasi Data Tidak Seimbang,” *METHOMIKA Jurnal Manajemen Informatika dan Komputerisasi Akuntansi*, vol. 4, no. 1, pp. 67–72, 2020.

[13] M. Sahami, D. David, E. Horvitz, and G. Building, “A Bayesian Approach to Filtering Junk E-Mail,” *Science*, pp. 55–61, 1998.

P-ISSN 2085-4315
E-ISSN 2502-8332

[14] A. Roihan, P. A. Sunarya, and A. S. Rafika, "Pemanfaatan Machine Learning dalam Berbagai Bidang," *(Indonesian Journal on Computer and Information Technology*, vol. 5, no. April, pp. 75–82, 2020.

[15] E. E. Amelia and I. Yustiana, "Analisis Sentimen Pada Ulasan Produk Pakaian UNIQLO Dengan Algoritma Naive Bayes," *Jurnal Sains Komputer & Informatika (J-SAKTI)*, vol. 8, no. 1, pp. 668–674, 2024.

[16] K. Henrys, "Importance of web scraping in e-commerce and e-marketing," no. January, pp. 1–10, 2021.

[17] Dania Siregar, Faroh Ladayya, Naufal Zhafran Albaqi, and Bintang Mahesa Wardana, "Penerapan Metode Support Vector Machines (SVM) dan Metode Naïve Bayes Classifier (NBC) dalam Analisis Sentimen Publik terhadap Konsep Child-free di Media Sosial Twitter," *Jurnal Statistika dan Aplikasinya*, vol. 7, no. 1, pp. 93–104, 2023.

[18] D. Oktavia, Y. R. Ramadahan, and Minarto, "Analisis Sentimen Terhadap Penerapan Sistem E-Tilang Pada Media Sosial Twitter Menggunakan Algoritma Support Vector Machine (SVM)," *Kajian Ilmiah Informatika dan Komputer*, vol. 4, no. 1, pp. 407–417, 2023.

[19] A. Widyastuti, Y. Astuti, B. A. Wisesa, and Hengki, "Analisis Sentimen pada Ulasan Produk dengan SVM dan Word2Vec Sentiment Analysis on Product Reviews with SVM and Word2Vec," *JIKO (Jurnal Informatika dan Komputer)*, vol. 8, no. 1, pp. 173–185, 2024.

[20] C. Hutto and E. Gilbert, "VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text," *Proceedings of English International*, pp. 216–225, 2014.

[21] A. Fernández, S. García, F. Herrera, and N. V. Chawla, "SMOTE for Learning from Imbalanced Data: Progress and Challenges, Marking the 15-year Anniversary," *Journal of Artificial Intelligence Research*, vol. 61, pp. 863–905, 2018.

[22] Z. Yang, "Recent Deep Learning Techniques on Short Text Classification," in *Proceedings of the 2022 6th International Conference on Electronic Information Technology and Computer Engineering*, 2023, pp. 231–235.

[23] A. G. Ghifari, G. Y. Ananada, and K. Purwandari, "A Comparative Sentiment Analysis of Public Opinion on A Comparative Sentiment Analysis of Public Opinion on Indonesia ' s National Football Coach Using CRNN and SVM Indonesia ' s National Football Coach Using CRNN and SVM," *Procedia Computer Science*, vol. 269, pp. 1485–1493, 2025.

[24] T. A. Almeida, J. M. G. Hidalgo, and A. Yamakami, "Contributions to the study of SMS spam filtering: New collection and results," *DocEng 2011 - Proceedings of the 2011 ACM Symposium on Document Engineering*, pp. 259–262, 2011.