

# PENERAPAN ALGORITMA SUPPORT VECTOR MACHINE UNTUK ANALISIS SENTIMEN ULASAN PELANGGAN TOKO LIVIA CIREBON DI SHOPEE

<sup>1</sup>Syaeful Annas\*, <sup>2</sup>Nana Suarna, <sup>3</sup>Irfan Ali, <sup>4</sup>Heliyanti Susana

<sup>1</sup>Teknik Informatika STMIK IKMI Cirebon, <sup>2</sup>Teknik Informatika STMIK IKMI Cirebon, <sup>3</sup>Rekaya Perangkat Lunak STMIK IKMI Cirebon, <sup>4</sup>Teknik Informatika STMIK IKMI Cirebon  
Jl. Perjuangan No.10B, Karyamulya, Kec. Kesambi, Kota Cirebon, Jawa Barat 45135

<sup>1</sup>syaefulannas33@gmail.com, <sup>2</sup>st\_nana@yahoo.com, <sup>3</sup>irfanaali0.0@gmail.com,

<sup>4</sup>heliheli@gmail.com

\*) Penulis Korespondensi

## Abstrak

*Analisis sentimen adalah proses yang bertujuan untuk memahami opini pelanggan dengan mengklasifikasikan ulasan menjadi sentimen positif, netral, atau negatif. Penelitian ini bertujuan untuk mengembangkan model analisis sentimen berbasis algoritma Support Vector Machine (SVM) terhadap ulasan pelanggan Toko Livia Cirebon di platform Shopee. Pendekatan penelitian dilakukan secara kuantitatif, dengan tahapan meliputi pengumpulan data, pra-pemrosesan teks (cleansing, normalisasi slang, tokenisasi, penghapusan stopword, dan stemming), pelabelan menggunakan Inset Lexicon, transformasi data teks menjadi vektor numerik dengan metode TF-IDF, pelatihan model SVM, serta evaluasi performa menggunakan metrik akurasi, precision, recall, dan F1-score. Model yang dikembangkan mencapai akurasi sebesar 91% dengan performa terbaik pada sentimen positif (F1-score 95%), meskipun performa pada kategori netral dan negatif masih memerlukan peningkatan. Hasil penelitian ini menunjukkan bahwa algoritma SVM efektif untuk analisis sentimen dalam e-commerce, memberikan wawasan strategis bagi pemilik usaha untuk menyusun strategi pemasaran dan meningkatkan kualitas layanan.*

**Kata Kunci:** Analisis Sentimen, E-commerce, Support Vector Machine, Ulasan Pelanggan, TF-IDF

## Abstract

*Sentiment analysis aims to understand customer opinions by classifying reviews into positive, neutral, or negative sentiments. This study aims to develop a sentiment analysis model using the Support Vector Machine (SVM) algorithm on customer reviews of Toko Livia Cirebon on the Shopee platform. The research adopts a quantitative approach, comprising stages such as data collection, text preprocessing (cleansing, slang normalization, tokenization, stopword removal, and stemming), sentiment labeling using the Inset Lexicon, text-to-vector transformation via the TF-IDF method, model training using SVM, and performance evaluation based on accuracy, precision, recall, and F1-score. The developed model achieved an accuracy of 91%, with the best performance observed for positive sentiment classification (F1-score 95%), although improvements are needed for neutral and negative categories. These findings indicate that the SVM algorithm is effective for sentiment analysis in e-commerce, providing strategic insights for business owners to refine marketing strategies and enhance service quality.*

**Keywords:** E-commerce, Sentiment Analysis, Support Vector Machine, TF-IDF, Customer Reviews

## PENDAHULUAN

Analisis sentimen yang sering disebut sebagai *opinion mining*, telah menjadi aspek penting dalam memahami opini pelanggan di era digital [1]. Melalui analisis ini, data tekstual seperti ulasan pelanggan dapat diolah untuk mengidentifikasi sentimen positif, negatif, atau netral [2]. Ulasan pelanggan pada platform *e-commerce* seperti Shopee memiliki pengaruh signifikan terhadap keputusan pembelian konsumen lainnya. Toko Livia Cirebon yang bergerak di bidang penjualan mainan tradisional seperti topeng dan barongsai, memiliki peluang untuk memanfaatkan analisis sentimen guna mengoptimalkan strategi pemasaran, meningkatkan kualitas produk, serta memahami kebutuhan pasar dengan lebih akurat.

Proses analisis manual terhadap ratusan atau ribuan ulasan pelanggan memiliki berbagai kendala, seperti keterbatasan waktu, sumber daya, dan potensi bias dalam interpretasi [3]. Ulasan juga sering kali mengandung bahasa informal, singkatan, atau kesalahan penulisan yang menambah kompleksitas dalam analisis otomatis. Algoritma yang mampu menangani data dalam jumlah besar secara efisien, seperti Support Vector Machine, diperlukan untuk mengatasi tantangan ini. Algoritma tersebut telah terbukti berhasil dalam banyak penelitian analisis sentimen. [4].

Penelitian ini bertujuan untuk mengembangkan model analisis sentimen

berbasis algoritma *SVM* yang mampu mengklasifikasikan ulasan pelanggan Toko Livia Cirebon di Shopee menjadi sentimen positif, negatif, dan netral. Model ini diharapkan dapat memberikan wawasan strategis kepada pemilik toko dalam rangka menyusun strategi pemasaran yang lebih efektif dan meningkatkan inovasi produk berdasarkan persepsi pelanggan.

Penelitian terdahulu menunjukkan keberhasilan penerapan *SVM* dalam berbagai konteks analisis sentimen. Penelitian oleh Rissa Nurfitriana Handayani menganalisis sentimen pada ulasan produk di Tokopedia menggunakan algoritma *SVM* yang dioptimalkan dengan *Particle Swarm Optimization (PSO)*. Hasilnya menunjukkan bahwa metode *SVM* tanpa optimasi menghasilkan akurasi sebesar 83,33%, sedangkan dengan *PSO* akurasi meningkat menjadi 88,89% [5]. Penelitian lain oleh Tjut Awaliyah Zuraiyah menganalisis sentimen pada ulasan produk di platform *e-commerce* menggunakan algoritma *Naïve Bayes*, *Support Vector Machine (SVM)*, dan *Recurrent Neural Network (RNN)*. Penelitian ini juga mengatasi masalah ketidakseimbangan data dengan menggunakan teknik oversampling *SMOTE*. Hasilnya menunjukkan bahwa metode *SVM* mencapai akurasi 88% [6]. Penelitian lain yang dilakukan oleh Dhiva Mustikananda membandingkan algoritma *Naïve Bayes* dan *SVM* untuk analisis sentimen terhadap ulasan produk kosmetik Aster di *marketplace* Shopee.

Hasilnya menunjukkan bahwa metode SVM mencapai akurasi tertinggi sebesar 87,56%, sedangkan Naïve Bayes hanya mencapai akurasi 82,13% [7].

Penelitian ini berfokus pada penerapan algoritma *SVM* untuk menganalisis ulasan pelanggan Toko Livia Cirebon di Shopee. Berbeda dengan penelitian terdahulu, penelitian ini mengintegrasikan *SVM* dengan tantangan analisis data informal dalam sektor *e-commerce*, memberikan solusi yang relevan bagi pelaku usaha kecil.

Algoritma *Support Vector Machine* digunakan untuk mengklasifikasikan ulasan pelanggan Toko Livia Cirebon di Shopee secara efektif dan akurat. Memberikan wawasan yang bermanfaat bagi optimalisasi strategi pemasaran dan inovasi produk.

## **METODE PENELITIAN**

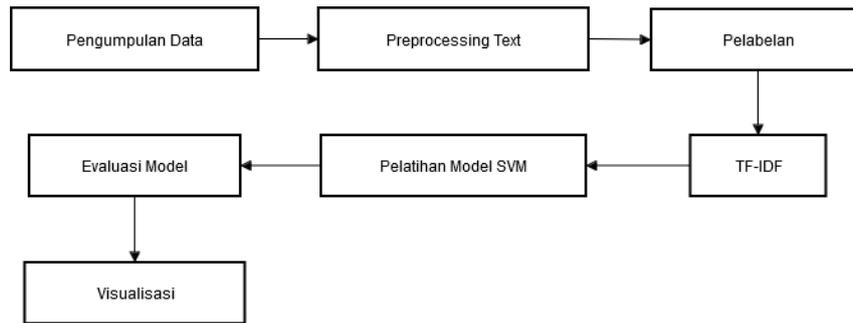
Penelitian ini berfokus pada penggunaan algoritma *SVM* untuk analisis umpan balik pelanggan. Proses penelitian dirancang secara sistematis untuk menjamin bahwa hasilnya akurat dan dapat diandalkan, melalui beberapa tahapan penting yang melibatkan pengumpulan data, pra-pemrosesan, pelabelan, transformasi data menjadi bentuk numerik, pelatihan model *SVM*, dan evaluasi performa model. Tujuan utama dari metode ini adalah mengembangkan model klasifikasi yang mampu

mengelompokkan ulasan pelanggan kedalam tiga kategori sentimen, yaitu positif, negatif, dan netral dengan akurasi tinggi. Setiap tahapan dalam metode penelitian di rancang untuk menangani utama dalam analisis data teks, seperti ketidakstabilan bahasa informal, variasi kata, serta kompleksitas tinggi dalam ulasan yang diambil dari platform *e-commerce*.

Gambar 1 menggambarkan alur metode penelitian analisis sentimen menggunakan *SVM*. Proses dimulai dengan pengumpulan data, dilanjutkan dengan *preprocessing* teks untuk pembersihan, pelabelan sentimen, dan transformasi fitur menggunakan TF-IDF. Data yang telah siap digunakan untuk pelatihan model *SVM*, diikuti evaluasi kinerja model, dan hasil akhirnya divisualisasikan untuk analisis lebih lanjut.

### **Pengumpulan Data**

Data ulasan pelanggan diambil secara manual dari platform Shopee, khususnya pada halaman Toko Livia Cirebon. Ulasan yang hanya berupa penilaian tanpa komentar tidak dimasukkan dalam penelitian. Data yang dikumpulkan berisi rentang waktu antara tahun 2020 hingga 2024 untuk memastikan relevansi dan aktualitas. Seluruh data yang telah dikumpulkan kemudian disimpan dalam bentuk *file Comma-Separated Values (CSV)* untuk memudahkan proses pengolahan lebih lanjut.



**Gambar 1. Metode Penelitian**

### ***Preprocessing Text***

Langkah berikutnya setelah pengumpulan data adalah melakukan preprocessing. *Preprocessing* diperlukan untuk meningkatkan kualitas dan akurasi hasil analisis sentimen [8]. Proses pra-pemrosesan data teks mencakup beberapa tahapan, seperti *cleansing*, slang *normalization*, *tokenizing*, *stopwords removal*, dan *stemming*. Berikut adalah penjelasan mengenai tahapan pra-pemrosesan data teks:

1. *Cleansing*

Tahap ini menghapus semua karakter non-alfabet dalam sebuah ulasan untuk mengurangi elemen yang tidak relevan dan tidak bermakna dalam analisis sentimen [9].

2. *Slang Normalization*

Tahap ini mengubah semua kata gaul dan disingkat menjadi bentuk standar menurut Kamus Besar Bahasa Indonesia (KBBI). Kata-kata tidak baku disesuaikan dengan kaidah bahasa[10].

3. *Tokenizing*

*Tokenizing* adalah proses memecah sebuah kalimat menjadi satu kata atau token [4].

4. *Stopword Removal*

*Stopword removal* adalah langkah untuk menghapus kata yang tidak memiliki kontribusi signifikan terhadap makna kalimat, seperti “di” atau “dengan” [11].

5. *Stemming*

*Stemming* adalah proses mengembalikan kata ke bentuk dasarnya dengan menghapus semua imbuhan yang melekat pada kata tersebut [12].

### **Pelabelan**

Pelabelan dilakukan menggunakan kamus *Inset Lexicon*. Kamus ini berisi 6.609 kata positif dan 3.609 kata negatif yang menggunakan bahasa indonesia. Setiap kata dilengkapi dengan bobot atau skor polaritas berkisar antara -5 hingga +5, yang mencerminkan tingkat intensitas sentimen dari kata tersebut [13].

## TF-IDF

Metode pembobotan *TF-IDF* (*Term Frequency-Inverse Document Frequency*) menilai setiap kata atau fitur dalam data teks untuk mengubahnya menjadi data numerik. TF-IDF merupakan ukuran statistik yang digunakan untuk menilai tingkat kepentingan sebuah kata dalam suatu dokumen [14]. Setelah transformasi ini, setiap ulasan memiliki representasi vektoral yang menunjukkan seberapa penting kata-kata yang ditemukan dalam dokumen tersebut.

## Pelatihan Model SVM

Proses pelatihan model adalah langkah krusial dalam *machine learning*. Tahap ini, model dilatih menggunakan data berlabel untuk memahami pola serta hubungan dalam data, sehingga dapat menjalankan tugas tertentu, seperti mengklasifikasikan sentimen.

## Evaluasi Model

Proses mengevaluasi kinerja model dengan data uji dikenal sebagai evaluasi model. Beberapa metrik yang digunakan untuk menilai performa model SVM adalah sebagai berikut:

### 1. Akurasi

Akurasi mengukur persentase prediksi yang tepat dari keseluruhan prediksi yang dilakukan model [15]. Nilai akurasi dapat dihitung menggunakan persamaan pada Rumus 1.

$$\text{Akurasi} = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

Jumlah data yang berhasil diprediksi dengan benar disebut TP, sedangkan TN merujuk pada jumlah data negatif yang diprediksi dengan benar. FP adalah jumlah data negatif yang keliru diprediksi sebagai positif, sementara FN adalah jumlah data positif yang keliru diprediksi sebagai negatif.

### 2. Precision

*Precision* menghitung tingkat ketepatan model dalam memprediksi data positif [16]. Nilai *precision* dapat dihitung menggunakan persamaan pada Rumus 2.

$$\text{Precision} = \frac{TP}{TP+FP} \quad (2)$$

### 3. Recall

*Recall* mengukur sejauh mana model mampu mendeteksi data positif dari seluruh data positif yang tersedia [17]. Nilai *recall* dapat dihitung menggunakan persamaan pada Rumus 3.

$$\text{Recall} = \frac{TP}{TP+FN} \quad (3)$$

### 4. F1-score

*F1-score* adalah rata-rata harmonis antara *precision* dan *recall*, yang bertujuan untuk menyeimbangkan kedua metrik tersebut [18]. Nilai *f1-score* dapat dihitung menggunakan persamaan pada Rumus 4.

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

## Visualisasi

Proses visualisasi data dilakukan setelah seluruh tahapan sebelumnya diselesaikan untuk memberikan gambaran yang lebih jelas tentang hasil analisis. *Word cloud* digunakan untuk memvisualisasikan distribusi kata-kata dominan pada sentimen positif, negatif, dan netral. *Confusion matrix* diterapkan selanjutnya untuk mengevaluasi akurasi dan performa prediksi model secara rinci. Diagram lingkaran digunakan terakhir untuk merepresentasikan distribusi proporsi sentimen dalam dataset secara keseluruhan.

## HASIL DAN PEMBAHASAN

Hasil analisis yang diperoleh dari proses pengumpulan data, *preprocessing text*, pelabelan, TF-IDF, Pelatihan model SVM, Evaluasi model, dan Visualisasi akan dibahas secara rinci.

### Hasil Pengumpulan Data

Langkah pertama dalam penelitian ini adalah pengumpulan data, yang akan digunakan untuk menganalisis sentimen. Data yang diambil berupa data sekunder sebanyak 607 ulasan yang diperoleh secara manual dari halaman Toko Livia Cirebon di platform Shopee melalui tautan [https://shopee.co.id/barongsai\\_mainan\\_anak](https://shopee.co.id/barongsai_mainan_anak). Data kemudian disimpan dalam bentuk *Comma Separated Values (CSV)*. Tabel 1 menunjukkan hasil pengumpulan dataset.

Tabel 1 menunjukkan sebagian *dataset* ulasan pelanggan Toko Livia Cirebon yang dianalisis.

### Hasil Preprocessing Text

Proses *preprocessing* dalam penelitian ini melibatkan sejumlah langkah yang dirancang untuk memastikan kualitas data teks sebelum dianalisis. Langkah pertama adalah *cleansing*, yang bertujuan untuk menghapus elemen-elemen seperti tanda baca, angka, simbol, dan karakter non-alfabet yang tidak relevan, sehingga mengurangi noise dalam data. Slang *normalization* dilakukan selanjutnya untuk mengonversi kata-kata tidak baku atau gaul menjadi bentuk formal sesuai standar bahasa, sehingga variasi penulisan dapat diminimalkan.

Proses *tokenizing* dilakukan untuk memecah teks menjadi unit kata atau token, sehingga setiap kata dapat dianalisis secara individu setelah itu. Langkah selanjutnya adalah menghapus kata-kata yang dianggap tidak mempengaruhi analisis, seperti *stopwords*, agar dapat lebih difokuskan pada kata-kata yang lebih penting. Proses dilanjutkan dengan *stemming*, yaitu mengembalikan kata-kata ke bentuk dasarnya agar data tetap konsisten. Tabel 2 menunjukkan hasil proses *cleansing* data ulasan pelanggan. Proses ini menghilangkan elemen yang tidak relevan, seperti tanda baca dan huruf kapital, sehingga menghasilkan data yang lebih bersih dan siap untuk analisis sentimen.

**Tabel 1. Dataset Ulasan Toko Livia Cirebon**

No	Ulasan
1	Makasih ponakan suka
2	Produk Sudah Saya Terima Dengan Baik Sesuai Deskripsi Sesuai Pesanan Saya Anak pun Bermain Dengan Senang nya Thx Seller Amanah Thx Kurir
3	Produknya sudah diterima dan sesuai orderan. Makasi ya...

**Tabel 2. Hasil Proses *Cleansing***

No	Sebelum <i>Cleansing</i>	Sesudah <i>Cleansing</i>
1	Makasih ponakan suka	makasih ponakan suka
2	Produk Sudah Saya Terima Dengan Baik Sesuai Deskripsi Sesuai Pesanan Saya Anak pun Bermain Dengan Senang nya Thx Seller Amanah Thx Kurir	produk sudah saya terima dengan baik sesuai deskripsi sesuai pesanan saya anak pun bermain dengan senang nya thx seller amanah thx kurir
3	Produknya sudah diterima dan sesuai orderan. Makasi ya...	produknya sudah diterima dan sesuai orderan makasi ya

**Tabel 3. Hasil Proses Slang *Normalization***

No	Sebelum Slang <i>Normalization</i>	Sesudah Slang <i>Normalization</i>
1	makasih ponakan suka	terima kasih ponakan suka
2	produk sudah saya terima dengan baik sesuai deskripsi sesuai pesanan saya anak pun bermain dengan senang nya thx seller amanah thx kurir	produk sudah saya terima dengan baik sesuai deskripsi sesuai pesanan saya anak pun bermain dengan senang nya terima kasih seller amanah terima kasih kurir
3	produknya sudah diterima dan sesuai orderan makasi ya	produknya sudah diterima dan sesuai orderan terima kasih ya

**Tabel 4. Hasil Proses *Tokenizing***

No	Sebelum <i>Tokenizing</i>	Sesudah <i>Tokenizing</i>
1	terima kasih ponakan suka	['terima', 'kasih', 'ponakan', 'suka']
2	produk sudah saya terima dengan baik sesuai deskripsi sesuai pesanan saya anak pun bermain dengan senang nya terima kasih seller amanah terima kasih kurir	['produk', 'sudah', 'saya', 'terima', 'dengan', 'baik', 'sesuai', 'deskripsi', 'sesuai', 'pesanan', 'saya', 'anak', 'pun', 'bermain', 'dengan', 'senang', 'nya', 'terima', 'kasih', 'seller', 'amanah', 'terima', 'kasih', 'kurir']
3	produknya sudah diterima dan sesuai orderan terima kasih ya	['produknya', 'sudah', 'diterima', 'dan', 'sesuai', 'orderan', 'terima', 'kasih', 'ya']

Tabel 3 menunjukkan hasil proses slang *normalization* pada data ulasan pengguna. Proses ini bertujuan untuk mengubah kata-kata slang atau informal menjadi bentuk formal yang sesuai dengan tata bahasa Indonesia. Tabel 4 menggambarkan hasil proses tokenisasi pada data ulasan pengguna. Proses

ini bertujuan untuk memecah teks menjadi unit-unit kata atau token. Tabel 5 menyajikan hasil proses *stopwords removal* pada data ulasan pengguna. Proses ini bertujuan untuk menghilangkan kata-kata yang tidak memiliki makna signifikan dalam analisis, sehingga meningkatkan efisiensi pemrosesan teks.

**Tabel 5. Hasil Proses Stopword Removal**

No	Sebelum Stopword Removal	Sesudah Stopword Removal
1	['terima', 'kasih', 'ponakan', 'suka']	['terima', 'kasih', 'ponakan', 'suka']
2	['produk', 'sudah', 'saya', 'terima', 'dengan', 'baik', 'sesuai', 'deskripsi', 'sesuai', 'pesanan', 'saya', 'anak', 'pun', 'bermain', 'dengan', 'senang', 'nya', 'terima', 'kasih', 'seller', 'amanah', 'terima', 'kasih', 'kurir']	'produk', 'terima', 'sesuai', 'deskripsi', 'sesuai', 'pesanan', 'bermain', 'senang', 'terima', 'kasih', 'seller', 'amanah', 'terima', 'kasih', 'kurir']
3	['produknya', 'sudah', 'diterima', 'dan', 'sesuai', 'orderan', 'terima', 'kasih', 'ya']	['produknya', 'diterima', 'sesuai', 'orderan', 'terima', 'kasih']

**Tabel 6. Hasil Proses Stemming**

No	Sebelum Stemming	Sesudah Stemming
1	['terima', 'kasih', 'ponakan', 'suka']	terima kasih ponakan suka
2	['produk', 'sudah', 'saya', 'terima', 'dengan', 'baik', 'sesuai', 'deskripsi', 'sesuai', 'pesanan', 'saya', 'anak', 'pun', 'bermain', 'dengan', 'senang', 'nya', 'terima', 'kasih', 'seller', 'amanah', 'terima', 'kasih', 'kurir']	produk terima sesuai deskripsi sesuai pesan main senang terima kasih seller amanah terima kasih kurir
3	['produknya', 'sudah', 'diterima', 'dan', 'sesuai', 'orderan', 'terima', 'kasih', 'ya']	produk terima sesuai order terima kasih

**Tabel 7. Hasil Pelabelan Sentimen**

No	Ulasan	Skor Polaritas	Sentimen
1	terima kasih ponakan suka	29	positive
2	produk terima sesuai deskripsi sesuai pesan main senang terima kasih seller amanah terima kasih kurir	7	positive
3	produk terima sesuai order terima kasih	12	positive

Tabel 6 menunjukkan hasil proses *stemming* pada data ulasan pengguna. Proses ini bertujuan untuk mengubah kata-kata menjadi bentuk dasar.

### Pelabelan

Tujuan pelabelan adalah untuk menentukan apakah ulasan pelanggan positif, negatif, atau netral. Penentuan sentimen dilakukan dengan menggunakan skor polaritas sebagai acuan. Aturan klasifikasi ditentukan berdasarkan nilai total skor polaritas, di mana ulasan dengan skor kurang dari 0 dianggap negatif, ulasan dengan skor lebih tinggi dari 0 dianggap positif, dan ulasan dengan skor sama dengan 0 dianggap netral [19].

Tabel 7 menyajikan hasil pelabelan sentimen berdasarkan skor polaritas yang diperoleh dari analisis ulasan pengguna. Ulasan pertama memiliki skor polaritas tertinggi sebesar 29, menunjukkan sentimen positif yang kuat. Ulasan kedua, dengan skor polaritas 7, tetap menunjukkan sentimen positif meskipun intensitasnya lebih rendah dibanding ulasan pertama.

### Hasil Pelatihan dan Evaluasi Model SVM

Tahap pelabelan selesai, kemudian data ulasan dengan label dikonversi ke dalam representasi numerik menggunakan metode *Term Frequency-Inverse Document Frequency (TF-IDF)*. Proses ini bertujuan

untuk memberikan bobot pada setiap kata dalam dokumen berdasarkan frekuensi kemunculannya serta relevansinya terhadap keseluruhan dataset. Transformasi *TF-IDF* menghasilkan vektor numerik yang dapat digunakan sebagai input untuk algoritma pembelajaran mesin [20].

Dataset yang telah diolah pada tahap pelatihan model dibagi menjadi dua subset, yaitu data uji 20% dan data latih 80%. Pembagian ini dilakukan untuk memastikan bahwa model dapat mempelajari pola dari data latih dan diuji kinerjanya pada data yang tidak terlihat sebelumnya. Algoritma *Support Vector Machine (SVM)* dengan *kernel* linear digunakan dalam proses pelatihan karena telah terbukti efektif untuk klasifikasi teks, terutama untuk data berukuran besar.

Evaluasi model dilakukan untuk mengukur kinerja algoritma dalam mengklasifikasikan sentimen ulasan ke dalam kategori positif, netral, dan negatif. Metode evaluasi menggunakan beberapa metrik standar, yaitu akurasi, *precision*, *recall*, dan *F1-score*. Tabel 8 menunjukkan performa

model *Support Vector Machine (SVM)* dalam mengklasifikasikan ulasan berdasarkan tiga kategori sentimen yaitu negatif, netral, dan positif. *Precision* pada kategori negatif mencapai 100%, yang menunjukkan seluruh prediksi kategori ini benar. *Recall* hanya 40%, yang berarti model hanya mampu mendeteksi 40% dari ulasan negatif yang sebenarnya. *F1-score* mencapai 57%, dengan jumlah data (*support*) sebanyak 5 ulasan. Kategori netral menunjukkan *precision* mencapai 100%, namun *recall* rendah di 11%, yang menandakan model hanya mengenali sebagian kecil ulasan netral. *F1-score* tercatat 20%, dengan jumlah data (*support*) sebanyak 9 ulasan. Kategori positif menunjukkan kinerja terbaik, dengan *precision* 91%, *recall* 100%, dan *F1-score* 95%. Hal ini mencerminkan kemampuan optimal model dalam mengidentifikasi seluruh ulasan positif dari total 108 ulasan. Tabel 9 menyajikan evaluasi keseluruhan kinerja model *Support Vector Machine (SVM)* dalam mengklasifikasikan sentimen ulasan menggunakan tiga metrik utama, yaitu *precision*, *recall*, dan *F1-score*.

**Tabel 8. Hasil Evaluasi Model SVM Berdasarkan Kategori Sentimen**

Kategori Sentimen	Precision	Recall	F1-Score	Support
Negatif	1.00	0.40	0.57	5
Netral	1.00	0.11	0.20	9
Positif	0.91	1.00	0.95	108

**Tabel 9. Hasil Evaluasi Model SVM Secara Keseluruhan**

Metode Evaluasi	Precision	Recall	F1-Score	Support
<i>Accuracy</i>			0.91	122
<i>Macro Avg</i>	0.97	0.50	0.57	122
<i>Weighted Avg</i>	0.92	0.91	0.88	122

Secara umum, model mencapai akurasi sebesar 91%, yang berarti 91% dari total 122 ulasan berhasil diklasifikasikan dengan benar. Nilai *macro average* menunjukkan rata-rata tidak berbobot dari semua kategori sentimen, dengan *precision* sebesar 97%, *recall* 50%, dan *F1-score* 57%. *Recall* yang rendah pada *macro average* mengindikasikan bahwa model kurang mampu mendeteksi secara konsisten sentimen dalam kategori minoritas, yaitu negatif dan netral. *Weighted average* memberikan gambaran kinerja yang mempertimbangkan kontribusi dari setiap kategori berdasarkan jumlah data (*support*). Nilai *precision* pada *weighted average* adalah 92%, *recall* 91%, dan *F1-score* 88%, mencerminkan performa model yang cukup baik dalam menangani kategori positif, yang memiliki jumlah data dominan, sambil tetap memberikan kontribusi pada kategori lain. Hasil evaluasi ini menunjukkan bahwa meskipun model bekerja sangat baik untuk sentimen positif, terdapat kebutuhan untuk mengoptimalkan performa pada sentimen negatif dan netral guna meningkatkan keseimbangan klasifikasi keseluruhan. Penelitian ini menggunakan algoritma Support Vector Machine (SVM) untuk menganalisis sentimen ulasan pelanggan Toko Livia Cirebon di Shopee, yang berbeda dengan penelitian sebelumnya dari segi pendekatan preprocessing dan evaluasi hasil.

Penelitian [5] mengoptimalkan SVM dengan *Particle Swarm Optimization (PSO)* untuk meningkatkan akurasi hingga 88,89%,

sedangkan penelitian ini mencapai akurasi 91% tanpa optimasi tambahan. Penelitian ini juga tidak menggunakan teknik *oversampling* seperti SMOTE yang diterapkan oleh penelitian [6], melainkan mengandalkan distribusi asli dataset untuk mengevaluasi performa model. Hasilnya menunjukkan performa unggul pada sentimen positif *F1-score* 95%, tetapi kategori netral dan negatif memerlukan peningkatan, sehingga penelitian ini memberikan kontribusi unik dalam konteks toko lokal di platform *e-commerce*.

### Visualisasi

Visualisasi data digunakan untuk memberikan gambaran yang lebih jelas dan berguna mengenai hasil analisis sentimen ulasan pelanggan. Penelitian ini melakukan visualisasi dengan tiga metode utama, yaitu Confusion Matrix, Wordcloud, dan diagram lingkaran. Penelitian ini, visualisasi dilakukan dengan tiga metode utama, yaitu *Confusion Matrix*, *Wordcloud*, dan diagram lingkaran.

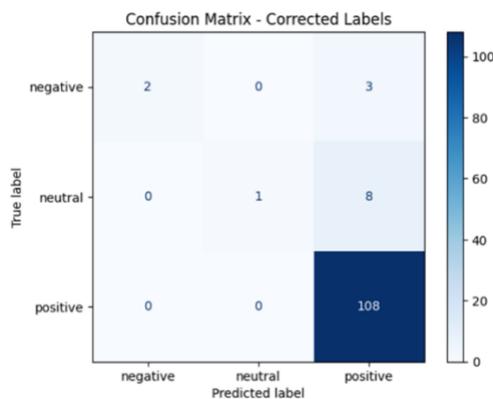
Gambar 2 menunjukkan *Confusion Matrix* yang menggambarkan hasil dari evaluasi model dalam mengklasifikasikan sentimen ulasan pelanggan. Model berhasil mengklasifikasikan sentimen positif dengan sangat baik, dengan 108 prediksi benar dan hanya sedikit kesalahan sebagai negatif (5) dan netral (2). Model mengidentifikasi 9 ulasan netral dengan benar, namun banyak yang salah dikategorikan sebagai positif (1). Sentimen negatif menunjukkan bahwa model berhasil memprediksi 5 ulasan dengan benar, namun

beberapa di antaranya salah diklasifikasikan sebagai positif (1). Secara keseluruhan, model menunjukkan kinerja terbaik dalam mendeteksi sentimen positif, namun masih membutuhkan perbaikan pada pengenalan sentimen netral dan negatif.

Gambar 3 menunjukkan visualisasi *word cloud* untuk sentimen positif, negatif, dan netral dari ulasan pelanggan Toko Livia Cirebon di Shopee. *Word cloud* untuk sentimen positif menunjukkan bahwa kata-kata seperti "harga," "bagus," dan "barang" mendominasi, mencerminkan apresiasi pelanggan terhadap harga produk yang kompetitif dan kualitas barang yang

memuaskan. *Word cloud* untuk sentimen negatif menyoroti kata-kata seperti "penyok," "mata," dan "rontok," yang mengindikasikan keluhan terkait kondisi fisik barang atau masalah kualitas.

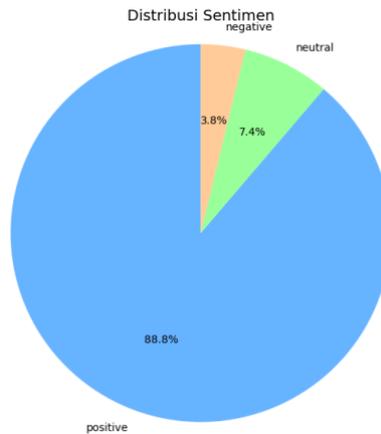
*Word cloud* untuk sentimen netral menampilkan kata-kata seperti "warna," "bahan," dan "barang," yang sering muncul, mencerminkan ulasan yang lebih deskriptif dan tidak cenderung ke arah tertentu. Analisis ini dapat membantu toko mengidentifikasi aspek yang perlu dipertahankan serta memperbaiki keluhan pelanggan untuk meningkatkan pengalaman belanja secara keseluruhan.



Gambar 2. Visualisasi *Confusion Matrix*



Gambar 3. Visualisasi *Wordcloud*



**Gambar 4. Visualisasi Distribusi Sentimen**

Gambar 4 menunjukkan distribusi sentimen ulasan pelanggan Toko Livia Cirebon di platform Shopee. Diagram lingkaran ini mengilustrasikan bahwa sebagian besar ulasan memiliki sentimen positif dengan persentase sebesar 88,8%, diikuti oleh sentimen netral sebanyak 7,4%, dan sentimen negatif yang hanya mencapai 3,8%. Hal ini menunjukkan bahwa mayoritas pelanggan memiliki pengalaman yang memuaskan terhadap layanan toko tersebut, sementara ulasan dengan sentimen negatif berada pada tingkat yang sangat rendah. Persentase ulasan negatif tetap perlu diperhatikan karena dapat menjadi indikator adanya aspek layanan yang membutuhkan perbaikan, sehingga pengalaman pelanggan dapat ditingkatkan secara menyeluruh.

## KESIMPULAN DAN SARAN

Penelitian ini berhasil mengimplementasikan algoritma *Support Vector Machine (SVM)* yang digunakan untuk mengevaluasi persepsi ulasan pelanggan Toko

Livia Cirebon di platform Shopee. Hasil analisis menunjukkan bahwa model SVM dengan kernel linear mampu mengklasifikasikan ulasan pelanggan ke dalam tiga kategori sentimen, yaitu positif, netral, dan negatif, dengan akurasi keseluruhan sebesar 91%. Model menunjukkan kinerja optimal pada ulasan dengan sentimen positif, dengan nilai *precision* sebesar 91%, *recall* 100%, dan *F1-score* 95%. Performa pada sentimen netral dan negatif masih memerlukan peningkatan, karena nilai recall untuk kategori tersebut relatif rendah. Visualisasi berupa *word cloud* dan diagram lingkaran memberikan wawasan tambahan tentang distribusi kata kunci dan proporsi sentimen dalam dataset.

Hasil penelitian ini menunjukkan bahwa algoritma SVM merupakan metode yang efektif untuk analisis sentimen ulasan pelanggan di *e-commerce*. Toko Livia Cirebon sebaiknya mempertahankan aspek-aspek yang telah dihargai pelanggan seperti harga produk yang kompetitif dan kualitas produk yang memuaskan. Keluhan terkait kondisi produk

yang rusak atau kualitas yang tidak sesuai dapat diperbaiki melalui peningkatan kontrol kualitas dan kebijakan pengembalian. Ulasan netral dapat dimanfaatkan untuk menggali kebutuhan pelanggan lebih dalam dengan menghubungi pihak yang memberi ulasan guna mendapatkan umpan balik tambahan.

Penelitian lebih lanjut dapat mencakup beberapa pengembangan untuk menyempurnakan hasil yang telah diperoleh. Pertama, metode pembobotan TF-IDF yang digunakan dalam penelitian ini dapat diperluas dengan eksplorasi teknik lain, seperti *word embedding* menggunakan algoritma *Word2Vec* atau *GloVe*, untuk meningkatkan representasi vektoral dari data teks. Kedua, pengembangan model dapat mencakup penggunaan algoritma *machine learning*, seperti *Random Forest* atau *Gradient Boosting*, sebagai pembanding terhadap performa SVM. Ketiga, untuk mengatasi tantangan dalam mendeteksi sentimen netral dan negatif, penelitian mendatang dapat memanfaatkan pendekatan *ensemble learning* atau pengolahan data yang lebih ekstensif, seperti augmentasi data.

#### DAFTAR PUSTAKA

- [1] N. Fitriyah, B. Warsito, and D. A. I. Maruddani, "Analisis Sentimen Gojek Pada Media Sosial Twitter Dengan Klasifikasi Support Vector Machine (Svm)," *J. Gaussian*, vol. 9, no. 3, pp. 376–390, 2020, doi:10.14710/j.gauss.v9i3.28932.
- [2] R. Parlika, S. Ilham Pradika, A. Muhammad Hakim, and K. Rachman N.M, "Analisis Sentimen Twitter Terhadap Bitcoin dan Cryptocurrency Berbasis Python TextBlob," *J. Ilm. Teknol. Inf. dan Robot.*, vol. 2, no. 2, pp. 33–37, 2020, doi:10.33005/jifti.v2i2.22.
- [3] A. P. Nardilasari, A. L. Hananto, S. S. Hilabi, T. Tukino, and B. Priyatna, "Analisis Sentimen Calon Presiden 2024 Menggunakan Algoritma SVM Pada Media Sosial Twitter," *JOINTECS (Journal Inf. Technol. Comput. Sci.)*, vol. 8, no. 1, p. 11, 2023, doi:10.31328/jointecs.v8i1.4265.
- [4] V. Fitriyana, L. Hakim, D. Candra Rini Novitasari, and A. Hanif, "Analisis Sentimen Ulasan Aplikasi Jamsostek Mobile Menggunakan Metode Support Vector Machine Vina," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 9, no. 4, pp. 3436–3442, 2022, doi:10.35957/jatisi.v9i4.3586.
- [5] R. Nurfitriana Handayani, "Optimasi Algoritma Support Vector Machine Untuk Analisis Sentimen Pada Ulasan Produk Tokopedia Menggunakan Pso," *Media Inform.*, vol. 20, no. 2, pp. 97–108, 2021.
- [6] T. A. Zuraiyah, M. M. Mulyati, and G. H. F. Harahap, "Perbandingan Metode Naïve Bayes, Support Vector Machine

- Dan Recurrent Neural Network Pada Analisis Sentimen Ulasan Produk E-Commerce,” *Multitek Indones.*, vol. 17, no. 1, pp. 27–43, 2023, doi: 10.24269/mtkind.v17i1.7092.
- [7] D. Mustikananda, D. E. Ratnawati, and B. Rahayudi, “Perbandingan Algoritma Naïve Bayes dan Support Vector Machine untuk Analisis Sentimen terhadap Review Produk Aster Kosmetik Malang Marketplace Shopee,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 6, no. 7, pp. 3137–3144, 2022.
- [8] A. I. Tanggraeni and M. N. N. Sitokdana, “Analisis Sentimen Aplikasi E-Government pada Google Play Menggunakan Algoritma Naïve Bayes,” *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 9, no. 2, pp. 785–795, 2022, doi: 10.35957/jatisi.v9i2.1835.
- [9] H. Tuhuteru, “Analisis Sentimen Masyarakat Terhadap Pembatasan Sosial Berksala Besar Menggunakan Algoritma Support Vector Machine,” *Inf. Syst. Dev.*, vol. 5, no. 2, pp. 7–13, 2020.
- [10] O. I. Gifari, M. Adha, F. Freddy, and F. F. S. Durrand, “Analisis Sentimen Review Film Menggunakan TF-IDF dan Support Vector Machine,” *J. Inf. Technol.*, vol. 2, no. 1, pp. 36–40, 2022, doi: 10.46229/jifotech.v2i1.330.
- [11] D. Oktavia, Y. R. Ramadahan, and Minarto, “Analisis Sentimen Terhadap Penerapan Sistem E-Tilang Pada Media Sosial Twitter Menggunakan Algoritma Support Vector Machine (SVM),” *KLIK Kaji. Ilm. Inform. dan Komput.*, vol. 4, no. 1, pp. 407–417, 2023, doi: 10.30865/klik.v4i1.1040.
- [12] F. F. Irfani, M. Triyanto, A. Dwi Hartanto, and Kusnawi, “Analisis Sentimen Review Aplikasi Ruangguru Menggunakan Algoritma Support Vector Machine,” *JBMI (Jurnal Bisnis, Manajemen, dan Inform.)*, vol. 16, no. 3, pp. 258–266, 2020, doi: 10.26487/jbmi.v16i3.8607.
- [13] D. Musfiroh, U. Khaira, P. E. P. Utomo, and T. Suratno, “Analisis Sentimen terhadap Perkuliahan Daring di Indonesia dari Twitter Dataset Menggunakan InSet Lexicon,” *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 1, no. 1, pp. 24–33, 2021, doi: 10.57152/malcom.v1i1.20.
- [14] J. A. Septian, T. M. Fachrudin, and A. Nugroho, “Analisis Sentimen Pengguna Twitter Terhadap Polemik Persepakbolaan Indonesia Menggunakan Pembobotan TF-IDF dan K-Nearest Neighbor,” *J. Intell. Syst. Comput.*, vol. 1, no. 1, pp. 43–49, 2019, doi: 10.52985/insyst.v1i1.36.
- [15] A. Pangestu, Y. Tajul Arifin, and R. Ade Safitri, “Analisis Sentimen Review Publik Pengguna Game Online Pada Platform Steam Menggunakan Algoritma Naïve Bayes,” *JATI (Jurnal*

- Mhs. Tek. Inform.*, vol. 7, no. 6, pp. 3106–3113, 2024, doi: 10.36040/jati.v7i6.8829.
- [16] E. R. Subhiyakto, Y. P. Astuti, N. Alexander, and E. Kartikadarma, “Analisis Sentimen Menggunakan Metode Naïve Bayes Untuk Mengetahui Respon Masyarakat Terhadap Vaksinasi,” *J. Ilm. Intech Inf. Technol. J. UMUS*, vol. 4, no. 02, pp. 179–188, 2022, doi: 10.46772/intech.v4i02.864.
- [17] T. Arifqi, N. Suarna, and W. Prihartono, “Penggunaan Naive Bayes Dalam Menganalisis Sentimen Ulasan Aplikasi Mcdonald’S Di Indonesia,” *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 8, no. 2, pp. 1949–1956, 2024, doi: 10.36040/jati.v8i2.8740.
- [18] Haris and D. Eka Ratnawati, “Analisis Sentimen berbasis Aspek terhadap Data Ulasan menggunakan Metode K-Nearest Neighbor (Studi Kasus: Aplikasi Olsera POS),” *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 7, no. 6, pp. 3041–3046, 2023, [Online]. Available: <http://j-ptiik.ub.ac.id>
- [19] M. Bagas Dikal Putra and E. Setiawan, “METODE LEXICON BASED UNTUK ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP KINERJA ISP,” *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 8, no. 6, pp. 12079–12087, 2024.
- [20] T. Safitri, Y. Umaidah, and I. Maulana, “Analisis Sentimen Pengguna Twitter Terhadap Grup Musik BTS Menggunakan Algoritma Support Vector Machine,” *J. Appl. Informatics Comput.*, vol. 7, no. 1, pp. 28–35, 2023, doi: 10.30871/jaic.v7i1.5039.