

# Optimasi Analisis Sentimen Ulasan *Sunscreen* di *E-Commerce* Menggunakan Algoritma SVM dan SMOTE

Ayi Andini<sup>1\*</sup>, Nining Rahaningsih<sup>1</sup>, Raditya Damar Dana<sup>1</sup>, Cep Lukman Rohmat<sup>1</sup>

<sup>1</sup>Program Studi Teknik Informatika, STMIK IKMI Cirebon, Indonesia

\*Email: ayiandini49@gmail.com

| Info Artikel   | Abstrak  |
|--|--|
| <p><b>Kata Kunci :</b><br/>                     analisis sentimen, algoritma <i>support vector machine</i> (SVM), ulasan, SMOTE</p> <p><b>Keywords :</b><br/>                     sentiment analysis, support vector machine (SVM) algorithm, reviews, SMOTE</p> | <p>Analisis sentimen terhadap ulasan pengguna di <i>e-commerce</i> membantu produsen memahami kepuasan pelanggan. Penelitian ini bertujuan untuk menganalisis sentimen ulasan produk <i>sunscreen</i> di <i>Facetology Official Shop</i> menggunakan algoritma <i>Support Vector Machine</i> (SVM). Data ulasan dikumpulkan melalui <i>scraping</i>, diberi label secara manual, dan diproses menggunakan metode <i>preprocessing</i> seperti <i>data cleaning</i>, <i>Case Folding</i>, <i>tokenizing</i>, <i>stopword removal</i>, serta SMOTE untuk menyeimbangkan data. Ekstraksi fitur dilakukan dengan TF-IDF, dan SVM digunakan untuk mengklasifikasikan sentimen menjadi positif, negatif, dan netral. Hasil penelitian menunjukkan model SVM dengan <i>kernel linear</i> mencapai akurasi 93%, presisi keseluruhan 95%, <i>recall</i> 91%, dan <i>F1-Score</i> 93%. Pendekatan ini menunjukkan peningkatan performa model dengan akurasi 93% setelah penerapan SMOTE untuk penyeimbangan data. Sentimen mayoritas positif, mengindikasikan tingkat kepuasan tinggi, meskipun ada ulasan negatif terkait efek samping produk. Teknik <i>preprocessing</i> dan penyeimbangan data terbukti efektif dalam meningkatkan performa model. Pendekatan dapat diaplikasikan untuk analisis sentimen produk serupa guna mendukung pemahaman perusahaan terhadap konsumen.</p> |
| <p><b>Tanggal Artikel</b><br/>                     Dikirim : 11 Desember 2024<br/>                     Direvisi : 2 Januari 2025<br/>                     Diterima : 10 Februari 2025</p>  | <p><b>Abstract</b></p> <p><i>Sentiment analysis of user reviews on e-commerce platforms helps producers understand customer satisfaction. This study aims to analyze the sentiment of sunscreen product reviews in the Facetology Official Shop using the Support Vector Machine (SVM) algorithm. Review data were collected through scraping, manually labeled, and processed using preprocessing methods such as data cleaning, case folding, tokenizing, stopwords removal, and SMOTE to balance the data. Feature extraction was performed using TF-IDF, and SVM was used to classify sentiments into positive, negative, and neutral categories. The results show that the SVM model with a linear kernel achieved an accuracy of 93%, an overall precision of 95%, a recall of 91%, and an F1-Score of 93%. This approach demonstrated improved model performance, with 93% accuracy achieved after applying SMOTE for data balancing. The majority of sentiments were positive, indicating a high level of customer satisfaction, although some negative reviews mentioned side effects of the product. The preprocessing techniques and data balancing proved effective in enhancing the model's performance. This approach can be applied to sentiment analysis of similar products to support companies in better understanding their consumers.</i></p>           |

## 1. PENDAHULUAN

Perkembangan teknologi informasi telah berdampak pada berbagai aspek kehidupan, termasuk dunia bisnis dan Pendidikan. Salah satu metode penting dalam bidang ini adalah analisis sentimen, yang berfungsi untuk mengelompokkan teks berdasarkan sentimen positif, negatif, atau netral. Dalam industri kecantikan, ulasan pelanggan menjadi sumber data yang berharga bagi produsen untuk memahami pandangan konsumen serta meningkatkan kualitas produk. Selain itu, analisis sentimen juga mendukung perusahaan dalam merancang strategi pemasaran yang lebih optimal [1].

Penelitian ini berfokus pada ulasan produk *sunscreen* dari *Facetology Official Shop*, dengan tujuan mengevaluasi tingkat kepuasan pelanggan. Masalah utama yang diidentifikasi adalah adanya ketidakpuasan pelanggan yang tidak dapat dilihat hanya melalui data penjualan, sehingga diperlukan pendekatan mendalam menggunakan algoritma *Support Vector Machine* (SVM). Algoritma ini dipilih karena kemampuannya menangani data berdimensi tinggi dan variatif, serta menghasilkan prediksi yang akurat. Namun, tantangan utama adalah ketidakseimbangan data ulasan dan variasi gaya penulisan, yang dapat memengaruhi performa model [2].

Penelitian sebelumnya menunjukkan efektivitas SVM dalam analisis sentimen. Pada ulasan *skincare*, akurasi mencapai 94% dengan nilai *precision*, *recall*, dan *F1-Score* masing-masing melebihi 0,9 [3]. Dalam studi ulasan *cushion* di *Female Daily*, SVM terbukti efektif dengan akurasi 78%, *precision* 90%, dan *recall* 100% [4]. Pada analisis produk *moisturizer* di *Twitter*, SVM berhasil mencapai akurasi 98%, *precision* 97%, *recall* 93%, dan *F1-Score* 95% [5]. Selain itu, penggunaan SMOTE telah terbukti meningkatkan performa model analisis sentimen pada *dataset* yang tidak seimbang, sebagaimana dibuktikan dalam penelitian [6]. Ekstraksi fitur menggunakan TF-IDF juga menjadi metode andalan dalam berbagai studi analisis sentimen, karena mampu mempresentasikan teks secara efektif untuk meningkatkan akurasi model [6]. Hasil ini menunjukkan potensi besar algoritma SVM, didukung oleh penerapan SMOTE dan TF-IDF, dalam analisis sentimen produk kecantikan.

Meskipun demikian, terdapat celah pada aspek *preprocessing* data yang dapat dieksplorasi lebih lanjut. Penelitian ini bertujuan mengoptimalkan penerapan SVM dalam analisis ulasan *sunscreen*, dengan fokus pada langkah-langkah *preprocessing* seperti *cleaning* data, *tokenizing*, dan *stopword removal*, untuk meningkatkan akurasi prediksi sentimen [7]. Selain itu, penelitian ini menggunakan *Python* dengan pustaka seperti *Numpy* dan *Scikit-Learn* untuk mendukung pengolahan data [8].

Penelitian ini diharapkan tidak hanya memberikan kontribusi teoritis tetapi juga praktis. Dengan analisis yang lebih mendalam, perusahaan dapat memahami preferensi konsumen, meningkatkan kualitas produk, dan memperkuat strategi pemasaran [9]. Hasilnya diharapkan memberikan wawasan baru tentang bagaimana sentimen konsumen mencerminkan tingkat kepuasan terhadap produk *sunscreen*.

Pertanyaan penelitian yang menjadi fokus utama dalam penelitian ini dapat dirumuskan secara rinci sebagai berikut :

1. Bagaimana menerapkan algoritma *Support Vector Machine* (SVM) untuk analisis sentimen ulasan produk *sunscreen* pada toko *Facetology Official Shop*?
2. Bagaimana mengukur performa akurasi, presisi, *recall*, dan *F1-Score* dari hasil perhitungan menggunakan algoritma *Support Vector Machine* (SVM)?

## 2. METODOLOGI PENELITIAN

### 2.1 Metode Penelitian

Penelitian ini menggunakan metode kuantitatif eksperimental dengan algoritma *Support Vector Machine* (SVM) untuk analisis sentimen ulasan pengguna produk *sunscreen*. Sentimen diklasifikasikan ke dalam kategori positif, negatif, dan netral melalui tahapan berikut:



Gambar 1 Tahapan Metode Penelitian

## 2.2 Sumber Data

Sumber data dalam penelitian ini berupa data sekunder yang diperoleh melalui teknik scraping menggunakan ekstensi Web Scraper pada Google Chrome dari ulasan produk *sunscreen* di *platform* e-commerce "Facetology Official Shop," serta dari jurnal ilmiah dan artikel yang membahas teori dan penelitian terkait topik serupa.

## 2.3 Teknik Pengumpulan Data

Data dalam penelitian ini dikumpulkan menggunakan teknik web scraping dengan ekstensi Web Scraper di Google Chrome. Setelah mengunduh dan memasang ekstensi, halaman web yang relevan dibuka, dan peta situs dibuat. Pemilih data ditentukan untuk memilih elemen yang akan diekstrak, kemudian proses scraping dijalankan. Data yang terkumpul disimpan dalam format CSV atau Excel untuk dianalisis.

## 2.4 Teknik Analisis Data

Teknik analisis data dalam penelitian ini melibatkan beberapa tahapan utama:

1. Pengumpulan Data  
Ulasan produk *sunscreen* dikumpulkan dari situs web Facetology Official Shop menggunakan *web scraping* dengan ekstensi Web Scraper di Google Chrome.
2. Pelabelan Data  
Data yang terkumpul diberi label secara manual oleh Ibu Hermin, M.Pd., seorang guru Bahasa Indonesia, untuk analisis sentimen yang lebih akurat.
3. Pemrosesan Data  
Tahap ini mengubah data mentah menjadi format yang sesuai untuk analisis lebih lanjut  
Langkah-langkah yang dilakukan meliputi:
  - *Cleansing Data*, Menghilangkan *noise*, duplikasi, dan informasi yang tidak relevan untuk memastikan data lebih bersih dan konsisten.
  - *Case Folding*, Mengubah seluruh teks menjadi huruf kecil untuk mengurangi variasi data akibat perbedaan kapitalisasi [10].
  - *Tokenizing*, Memisahkan teks menjadi kata-kata terpisah agar dapat diolah pada level kata [11].
  - *Stopword removal*, Menghapus kata-kata seperti "dan", "dengan", dan "yang" yang tidak memiliki makna dalam analisis sentimen, berdasarkan daftar kata umum dalam bahasa Indonesia yang tidak relevan untuk analisis [12].
  - *Stemming*, Proses mengubah kata-kata menjadi bentuk dasar untuk menyederhanakan variasi kata, sehingga maknanya tetap konsisten dan lebih mudah dianalisis [12].
4. Ekstraksi Fitur  
TF-IDF digunakan untuk memberikan bobot pada istilah berdasarkan frekuensi, membantu model fokus pada kata-kata yang paling relevan [13].

5. Implementasi Algoritma (SVM)

Analisis sentimen dilakukan dengan menggunakan *Support Vector Machine* (SVM) pada data yang telah diproses sebelumnya [14]. Kernel linear dipilih karena efektif untuk data teks berdimensi tinggi seperti ulasan dan mampu memberikan hasil yang akurat dan efisien. Parameter **C (Regularization Parameter)** diatur untuk menjaga keseimbangan antara margin dan tingkat kesalahan klasifikasi [15]. Sementara kernel non-linear, seperti RBF, Polinomial, tidak digunakan karena data ini tidak memerlukan transformasi yang kompleks.

6. Evaluasi

Evaluasi dilakukan dengan menghitung akurasi, presisi, dan *recall* menggunakan *Confusion Matrix* untuk menilai kinerja model SVM [16].

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

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

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

$$F1 - Score = 2 \times \frac{Presisi \times Recall}{Presisi + Recall} \quad (4)$$

7. Visualisasi

Evaluasi dilakukan dengan menghitung akurasi, presisi, dan *recall* menggunakan *Confusion Matrix* untuk menilai performa model SVM [17].

3. Hasil dan Pembahasan

Proses analisis sentimen menggunakan algoritma SVM mencakup tahapan *preprocessing* data, ekstraksi fitur, penerapan SMOTE *Upsampling*, implementasi algoritma SVM, evaluasi model, serta visualisasi, yang diuraikan secara rinci sebagai berikut:

3.1 *Preprocessing* Data

Data yang telah dilabeli disimpan dalam format “xlsx” dan dipreproses menggunakan *Google Collaboratory* dengan *Python* untuk membersihkan *noise*, memperbaiki bahasa, dan menstrukturkan data. Tahapan *preprocessing* adalah sebagai berikut:

1. *Cleaning* Data

Tahap ini merupakan bagian penting dari *preprocessing* yang bertujuan untuk mengoreksi atau menghapus data yang tidak valid, tidak lengkap, atau tidak relevan, termasuk URL, tagar, angka, karakter khusus, dan lain-lain. Tujuan dari pembersihan data adalah untuk memastikan bahwa data yang akan dianalisis memiliki kualitas yang baik dan bebas dari elemen-elemen yang dapat mempengaruhi hasil analisis atau performa model.

Tabel 1 Hasil *Cleaning* Data

| Sebelum  | Label   | Sesudah  | Label   |
|--|---------|--|---------|
| <i>Sunscreen</i> ini memberikan perlindungan yang kuat tanpa menyebabkan kulit saya berkilau. Sangat nyaman digunakan sehari-hari dan tidak terasa berat di kulit. | Positif | <i>sunscreen</i> ini memberikan perlindungan yang kuat tanpa menyebabkan kulit saya berkilau sangat nyaman digunakan sehari-hari dan tidak terasa berat di kulit | Positif |
| Teksturnya baik, tapi tidak berpengaruh pada kulit saya  | Netral  | teksturnya baik tapi tidak berpengaruh pada kulit saya   | Netral  |
| Tidak ada efek samping, tapi juga tidak ada perubahan yang terlihat  | Netral  | tidak ada efek samping tapi juga tidak ada perubahan yang terlihat   | Netral  |

2. *Case Folding*

Pada langkah tahap ini, huruf kapital diubah menjadi huruf kecil menggunakan Python. Hal ini memastikan kata-kata seperti “Bagus” dan “bagus” diperlakukan sama, sehingga mengurangi variasi kata yang tidak perlu dan meningkatkan akurasi analisis teks.

**Tabel 2 Hasil Case Folding**

| Sebelum  | Sesudah  |
|--|--|
| <i>Sunscreen</i> ini memberikan perlindungan yang kuat tanpa menyebabkan kulit saya berkilau. Sangat nyaman digunakan sehari-hari dan tidak terasa berat di kulit. | <i>sunscreen</i> ini memberikan perlindungan yang kuat tanpa menyebabkan kulit saya berkilau sangat nyaman digunakan sehari-hari dan tidak terasa berat di kulit |
| Teksturnya baik, tapi tidak berpengaruh pada kulit saya  | teksturnya baik tapi tidak berpengaruh pada kulit saya   |
| Tidak ada efek samping, tapi juga tidak ada perubahan yang terlihat  | tidak ada efek samping tapi juga tidak ada perubahan yang terlihat   |

Pada Tabel 1, hasil dari proses *Case Folding* adalah perubahan semua huruf dalam teks menjadi huruf kecil.

3. *Tokenizing*

Pada tahap *tokenizing*, ulasan dipecah menjadi kata-kata atau “token” untuk menyederhanakan analisis, sehingga lebih mudah untuk mengidentifikasi istilah-istilah kunci dan meningkatkan akurasi pemrosesan data teks.

**Tabel 3 Hasil Tokenizing**

| Sebelum  | Sesudah  |
|--|--|
| <i>Sunscreen</i> ini memberikan perlindungan yang kuat tanpa menyebabkan kulit saya berkilau. Sangat nyaman digunakan sehari-hari dan tidak terasa berat di kulit. | [ <i>sunscreen</i> , ini, memberikan, perlindungan, yang, kuat, tanpa, menyebabkan, kulit, saya, berkilau, sangat, nyaman, digunakan, sehari-hari, dan, tidak, terasa, berat, di, kulit] |
| Teksturnya baik, tapi tidak berpengaruh pada kulit saya  | [teksturnya, baik, tapi, tidak, berpengaruh, pada, kulit, saya]  |
| Tidak ada efek samping, tapi juga tidak ada perubahan yang terlihat  | [tidak, ada, efek, samping, tapi, juga, tidak, ada, perubahan, yang, terlihat]   |

Pada Tabel 2, ulasan telah diuraikan menjadi beberapa kata, sehingga analisis kata per kata menjadi lebih mudah dan pola serta konteksnya dapat dipahami dengan lebih jelas.

4. *Stopword removal* dan *Stemmer*

Tahap *stopword removal* dan *stemming* mengubah kata menjadi bentuk dasarnya. *stopword removal* dan tahap *stemming* mengubah kata menjadi bentuk dasarnya. *stopword removal* menghilangkan kata-kata yang umum dan tidak relevan, sementara *stemming* mempertimbangkan imbuhan untuk menghasilkan bentuk dasar yang lebih akurat dan bermakna.

| cleaned_text                                      | label   | Review_len | punct | tokens   | removed_stopwords                                  | stemmer_review                                     |
|---|---------|------------|-------|--|--|--|
| sunscreen ini memberikan perlindungan yang kua... | Positif | 135        | 2.2   | [sunscreen, ini, memberikan, perlindungan, yan...  | sunscreen memberikan perlindungan kuat menyeba...  | sunscreen beri lindung kuat sebab kulit kilau ...  |
| teksturnya baik tapi tidak berpengaruh pada ku... | Netral  | 48         | 2.1   | [teksturnya, baik, tapi, tidak, berpengaruh, p...  | teksturnya baik tidak berpengaruh kulit            | tekstur baik tidak pengaruh kulit                  |
| tidak ada efek samping tapi juga tidak ada per... | Netral  | 57         | 1.8   | [tidak, ada, efek, samping, tapi, juga, tidak, ... | ada efek samping juga ada perubahan terlihat       | ada efek samping juga ada ubah lihat               |
| cukup oke tapi tidak ada yang istimewa seperti... | Netral  | 62         | 3.2   | [cukup, oke, tapi, tidak, ada, yang, istimewa, ... | cukup oke tidak yang istimewa yang orang-orang ... | cukup oke tidak yang istimewa yang orang-orang ... |
| ss ini ringan tapi hasilnya biasa saja            | Netral  | 33         | 3.0   | [ss, ini, ringan, tapi, hasilnya, biasa, saja]     | ss ringan hasilnya biasa                           | ss ringan hasil biasa                              |

**Gambar 2 Hasil Stopword removal dan Stemmer**



```
+ Code + Text All changes saved Connect + Gemini [user] [settings] [refresh] [close]
Penerapan SMOTE untuk menyeimbangkan Data
[] # Penerapan SMOTE untuk menyeimbangkan data
from imblearn.over_sampling import SMOTE
smote = SMOTE(random_state=42)
X_train_balanced, y_train_balanced = smote.fit_resample(X_train_vect, y_train)
```

Gambar 4 Kode Program Untuk Proses SMOTE Upsampling

### 3.4 Implementasi Algoritma (SVM)

Proses implementasi ini meliputi tahap pelatihan model *Support Vector Machine* (SVM). Langkah pertama dalam proses pelatihan SVM dimulai dengan inisialisasi model SVM menggunakan kernel tertentu.

```
▼ Pelatihan Model SVM
# Pelatihan Model SVM
from sklearn.svm import SVC
classifier = SVC(kernel='linear', random_state=0)
classifier.fit(X_train_balanced, y_train_balanced)
svm_pred = classifier.predict(X_test_vect)
```

Gambar 5 Kode Program Untuk Proses Implementasi SVM

Gambar 6 menunjukkan bahwa model SVM dengan Linear Kernel memberikan hasil terbaik untuk mengklasifikasikan data penelitian.

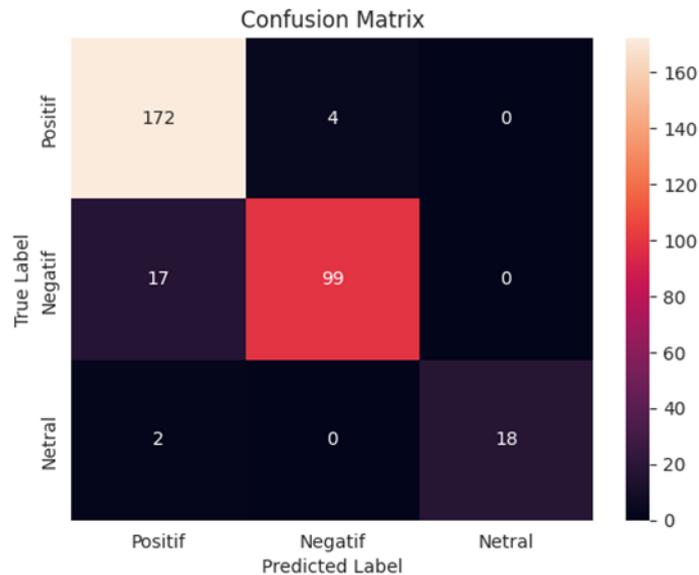
### 3.5 Evaluasi Model

Tahap evaluasi model bertujuan untuk mengukur sejauh mana model mampu membuat prediksi dengan akurat.

|              | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| Negatif      | 0.96      | 0.85   | 0.90     | 116     |
| Netral       | 1.00      | 0.90   | 0.95     | 20      |
| Positif      | 0.90      | 0.98   | 0.94     | 176     |
| accuracy     |           |        | 0.93     | 312     |
| macro avg    | 0.95      | 0.91   | 0.93     | 312     |
| weighted avg | 0.93      | 0.93   | 0.93     | 312     |

Gambar 6 Hasil Evaluasi Model

Hasil evaluasi menunjukkan akurasi sebesar 93%, seperti yang ditunjukkan pada Gambar 7.



Gambar 7 Hasil *Confusion Matrix*

Pada Gambar 8, hasil klasifikasi model SVM ditampilkan untuk tiga kelas sentimen: "Positif," "Negatif," dan "Netral." *Confusion Matrix* ini memberikan gambaran tentang prediksi yang benar dan salah untuk setiap kelas, yang digunakan untuk menghitung akurasi, presisi, *recall*, dan *F1-score*.

### 1. Penjelasan tentang *Confusion Matrix*

- Kelas Positif
  - *True Positif* (TP) = 172 (Prediksi benar Positif)
  - *False Positif* (FP) = 17+2 = 19 (Jumlah sampel kelas Negatif dan Netral yang salah prediksi sebagai Positif)
  - *False Negatif* (FN) = 4+0 = 4 (Jumlah sampel kelas Positif yang salah diprediksi sebagai Negatif atau Netral)
- Kelas Negatif
  - *True Positif* (TP) = 99 (Prediksi benar Negatif)
  - *False Positif* (FP) = 4+0 = 4 (Jumlah sampel kelas Positif dan Netral yang salah diprediksi sebagai Negatif)
  - *False Negatif* (FN) = 17+0 = 17 (Jumlah sampel kelas Negatif yang salah diprediksi sebagai Positif atau Netral)
- Kelas Netral
  - *True Positif* (TP) = 18 (Prediksi benar Netral)
  - *False Positif* (FP) = 0+0 = 0 (Tidak ada sampel dari kelas Positif atau Negatif yang salah diprediksi sebagai Netral)
  - *False Negatif* (FN) = 2+0 = 2 (Jumlah sampel kelas Netral yang salah diprediksi sebagai Positif dan Negatif)

### 2. Perhitungan Matrix

#### 1. Accuracy

$$\begin{aligned}
 Accuracy &= \frac{TP + TN}{TP + TN + FP + FN} \\
 &= \frac{172 + 99 + 18}{172 + 4 + 0 + 17 + 99 + 0 + 2 + 0 + 18} \\
 &= \frac{289}{312} \approx 0.93\% \text{ atau } 93\%
 \end{aligned}$$

#### 2. Presisi

$$Presisi = \frac{TP}{TP + FP}$$

- Presisi Kelas Positif

$$\begin{aligned} \text{Presisi Positif} &= \frac{TP_{\text{positif}}}{TP_{\text{positif}} + FP_{\text{positif}}} \\ &= \frac{172}{172 + 19} = \frac{172}{191} \approx 0.90\% \text{ atau } 90\% \end{aligned}$$

- Presisi Kelas Negatif

$$\begin{aligned} \text{Presisi Negatif} &= \frac{TP_{\text{negatif}}}{TP_{\text{negatif}} + FP_{\text{negatif}}} \\ &= \frac{99}{99 + 4} = \frac{99}{103} \approx 0.96\% \text{ atau } 96\% \end{aligned}$$

- Presisi Kelas Netral

$$\begin{aligned} \text{Presisi Netral} &= \frac{TP_{\text{netral}}}{TP_{\text{netral}} + FN_{\text{netral}}} \\ &= \frac{18}{18 + 0} = \frac{18}{18} = 1.00\% \text{ atau } 100\% \end{aligned}$$

- Presisi Keseluruhan

$$\begin{aligned} \text{Presisi Keseluruhan} &= \frac{\text{Presisi Positif} + \text{Presisi Negatif} + \text{Presisi Netral}}{3} \\ &= \frac{90\% + 96\% + 100\%}{3} = \frac{286}{3} \approx 95\% \end{aligned}$$

### 3. Recall

$$\text{Recall} = \frac{TP}{TP + FN}$$

- Recall Kelas Positif

$$\begin{aligned} \text{Recall Positif} &= \frac{TP_{\text{positif}}}{TP_{\text{positif}} + FN_{\text{positif}}} \\ &= \frac{172}{172 + 4} = \frac{172}{176} \approx 0.98\% \text{ atau } 98\% \end{aligned}$$

- Recall Kelas Negatif

$$\begin{aligned} \text{Recall Negatif} &= \frac{TP_{\text{negatif}}}{TP_{\text{negatif}} + FN_{\text{negatif}}} \\ &= \frac{99}{99 + 17} = \frac{99}{116} \approx 0.85\% \text{ atau } 85\% \end{aligned}$$

- Recall Kelas Netral

$$\begin{aligned} \text{Recall Netral} &= \frac{TP_{\text{netral}}}{TP_{\text{netral}} + FN_{\text{netral}}} \\ &= \frac{18}{18 + 2} = \frac{18}{20} = 0.90\% \text{ atau } 90\% \end{aligned}$$

- Recall Keseluruhan

$$\begin{aligned} \text{Recall Keseluruhan} &= \frac{\text{Recall Positif} + \text{Recall Negatif} + \text{Recall Netral}}{3} \\ &= \frac{98\% + 85\% + 90\%}{3} \approx 91\% \end{aligned}$$

#### 4. *F1-Score*

$$F1 - Score = 2 \times \frac{Presisi \times Recall}{Presisi + Recall}$$

➤ *F1-Score* Kelas Positif

$$\begin{aligned} F1 - Score &= 2 \times \frac{Presisi Positif \times Recall Positif}{Presisi Positif + Recall Positif} \\ &= 2 \times \frac{0.90 \times 0.98}{0.90 + 0.98} \approx 0.94\% \text{ atau } 94\% \end{aligned}$$

➤ *F1-Score* Kelas Negatif

$$\begin{aligned} F1 - Score &= 2 \times \frac{Presisi Negatif \times Recall Negatif}{Presisi Negatif + Recall Negatif} \\ &= 2 \times \frac{0.96 \times 0.85}{0.96 + 0.85} \approx 0.90\% \text{ atau } 90\% \end{aligned}$$

➤ *F1-Score* Kelas Netral

$$\begin{aligned} F1 - Score &= 2 \times \frac{Presisi Netral \times Recall Netral}{Presisi Netral + Recall Netral} \\ &= 2 \times \frac{1.00 \times 0.90}{1.00 + 0.90} \approx 0.95\% \text{ atau } 95\% \end{aligned}$$

➤ *F1-Score* Keseluruhan

$$\begin{aligned} F1 - Score &= \frac{F1 - score Positif + F1 - score Negatif + F1 - score Netral}{3} \\ &= \frac{94\% + 90\% + 95\%}{3} \approx 93\% \end{aligned}$$

Hasil penelitian menunjukkan bahwa model SVM dengan kernel linear mencapai akurasi 93%, presisi 90% untuk kelas positif, 96% untuk kelas negatif, dan 100% untuk kelas netral. Sementara itu, nilai *recall* tercatat sebesar 98% untuk kelas positif, 85% untuk kelas negatif, dan 90% untuk kelas netral. *F1-Score* keseluruhan mencapai 93%, menunjukkan keseimbangan antara presisi dan *recall* dalam mengklasifikasikan data di ketiga kelas sentimen. Jika dibandingkan dengan penelitian serupa, seperti penelitian [18] yang menerapkan SVM dengan kernel RBF, model ini menunjukkan keunggulan terutama dalam akurasi dan *F1-Score*, khususnya pada kelas netral yang sering menjadi tantangan dalam analisis sentimen. Temuan ini memperkuat hasil penelitian sebelumnya tentang keandalan SVM, sekaligus menunjukkan bahwa kernel linear mampu memberikan performa optimal pada dataset dengan struktur kompleks.

Meskipun model menunjukkan performa yang baik secara keseluruhan, tantangan terlihat pada kelas negatif yang memiliki *recall* lebih rendah. Hal ini kemungkinan disebabkan oleh variasi bahasa yang digunakan dalam ulasan negatif, seperti penggunaan sinonim, eufemisme, atau struktur kalimat yang kompleks, serta jumlah data ulasan negatif yang lebih sedikit dibandingkan dengan kelas lainnya. Temuan ini menunjukkan perlunya perbaikan, seperti penambahan data ulasan negatif atau penerapan teknik *preprocessing* yang lebih canggih, untuk meningkatkan kemampuan model dalam mendeteksi sentimen negatif. Meskipun demikian, hasil penelitian ini tetap menunjukkan potensi model SVM dalam analisis sentimen pada *platform e-commerce*.

### 3.6 Visualisasi

#### 1. Visualisasi dengan Diagram *Pie*

Diagram lingkaran menampilkan proporsi setiap kategori sentimen (positif, negatif, dan netral), sehingga lebih mudah untuk membandingkan jumlah ulasan di setiap kategori dan memahami kepuasan pengguna.





## DAFTAR PUSTAKA

- [1] F. Pradana Rachman, H. Santoso, and A. History, "Jurnal Teknologi dan Manajemen Informatika Perbandingan Model Deep Learning untuk Klasifikasi Sentiment Analysis dengan Teknik Natural Language Processing Article Info ABSTRACT," *J. Teknol. dan Manaj. Inform.*, vol. 7, no. 2, pp. 103–112, 2021, [Online]. Available: <http://http/jurnal.unmer.ac.id/index.php/jtmi>
- [2] I. S. K. Idris, Y. A. Mustofa, and I. A. Salihi, "Analisis Sentimen Terhadap Penggunaan Aplikasi Shopee Menggunakan Algoritma Support Vector Machine (SVM)," *Jambura J. Electr. Electron. Eng.*, vol. 5, no. 1, pp. 32–35, 2023, doi: 10.37905/jjee.v5i1.16830.
- [3] Harnelia and R. Adi Saputra, "Analisis Sentimen Review Skincare Skintific Dengan Algoritma Support Vector Machine (Svm)," *J. Inform. dan Tek. Elektro Terap.*, vol. 12, no. 2, 2024, doi: 10.23960/jitet.v12i2.4095.
- [4] S. Ariqoh, M. A. Sunandar, and Y. Muhyidin, "Analisis Sentimen Pada Produk Cushion Di Website Female Daily Menggunakan Metode Support Vector Machine (Svm)," *STORAGE J. Ilm. Tek. dan Ilmu Komput.*, vol. 2, no. 3, pp. 137–142, 2023, doi: 10.55123/storage.v2i3.2345.
- [5] R. Cantika Larasati, C. Dewi, and C. H. Juli, "Analisis sentimen produk kecantikan jenis moisturizer di twitter menggunakan algoritma super vector machine," *TEKINKOM*, vol. 7, no. 1, pp. 124–134, 2024, doi: 10.37600/tekinkom.v7i1.1243.
- [6] I. G. B. A. Budaya and I. K. P. Suniantara, "Comparison of Sentiment Analysis Algorithms with SMOTE Oversampling and TF-IDF Implementation on Google Reviews for Public Health Centers," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 4, no. 3, pp. 1077–1086, 2024, doi: 10.57152/malcom.v4i3.1459.
- [7] M. I. Buana and D. B. Arianto, "Analisis Sentimen Ulasan Pengguna Aplikasi ZenPro dengan Implementasi Algoritma Support Vector Machine ( SVM )," *Adopsi Teknol. dan Sist. Inf.*, vol. 3, no. 1, pp. 45–52, 2024, doi: 10.30872/atasi.v3i1.1092.
- [8] N. Yolanda, I. H. Santi, and D. F. H. Permadi, "Analisis Sentimen Analisis Sentimen Popularitas Aplikasi Moodle dan Edmodo Menggunakan Algoritma Support Vector Machine," *J. Algoritm.*, vol. 3, no. 1, pp. 48–59, 2022, doi: 10.35957/algoritme.v3i1.3313.
- [9] M. Hamka, N. Alfatar, and D. Ratna Sari, "Analisis Sentimen Produk Kecantikan Jenis Serum Menggunakan Algoritma Naïve Bayes Classifier," *J. Sist. Komput. dan Inform.*, vol. 4, no. 1, p. 64, 2022, doi: 10.30865/json.v4i1.4740.
- [10] I. W. B. Suryawan, N. W. Utami, and K. Q. Fredlina, "Analisis Sentimen Review Wisatawan pada Objek Wisata Ubud Menggunakan Algoritma Support Vector Machine," *J. Inform. Teknol. dan Sains*, vol. 5, no. 1, pp. 133–140, 2023.
- [11] A. Karel Maulaya and Junadhi, "Analisis Sentimen Menggunakan Support Vector Machine Masyarakat Indonesia Di Twitter Terkait Bjorka," *J. CoSciTech (Computer Sci. Inf. Technol.*, vol. 3, no. 3, pp. 495–500, 2022, doi: 10.37859/coscitech.v3i3.4358.
- [12] G. Radiana and A. Nugroho, "Analisis Sentimen Berbasis Aspek Pada Ulasan Aplikasi Kai Access Menggunakan Metode Support Vector Machine," *J. Pendidik. Teknol. Inf.*, vol. 6, no. 1, pp. 1–10, 2023, doi: 10.37792/jukanti.v6i1.836.
- [13] A. D. Pratama and H. Hendry, "Analisa Sentimen Masyarakat Terhadap Penggunaan Chatgpt Menggunakan Metode Support Vector Machine (Svm)," *JUPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.*, vol. 9, no. 1, pp. 327–338, 2024, doi: 10.29100/jupi.v9i1.4285.
- [14] D. Angraina Fitri and A. Putri, "Analisis Sentimen Pengguna Aplikasi Google Meet Menggunakan Algoritma Support Vector Machine," *J. CoSciTech (Computer Sci. Inf. Technol.*, vol. 3, no. 3, pp. 472–478, 2022, doi: 10.37859/coscitech.v3i3.4260.
- [15] T. M. Permata Aulia, N. Arifin, and R. Mayasari, "Perbandingan Kernel Support Vector Machine (Svm) Dalam Penerapan Analisis Sentimen Vaksinasi Covid-19," *SINTECH (Science Inf. Technol. J.*, vol. 4, no. 2, pp. 139–145, 2021, doi: 10.31598/sintechjournal.v4i2.762.
- [16] 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.
- [17] J. Ipmawati, Saifulloh, and Kusnawi, "Analisis Sentimen Tempat Wisata Berdasarkan Ulasan pada Google Maps Menggunakan Algoritma Support Vector Machine," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 4, no. 1, pp. 247–256, 2024, doi: 10.57152/malcom.v4i1.1066.
- [18] S. Rabbani, D. Safitri, N. Rahmadhani, A. A. F. Sani, and M. K. Anam, "Perbandingan Evaluasi Kernel SVM untuk Klasifikasi Sentimen dalam Analisis Kenaikan Harga BBM," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 3, no. 2, pp. 153–160, 2023, doi: 10.57152/malcom.v3i2.897.