

ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP SKINCARE DENGAN METODE *SUPPORT VECTOR MACHINE* (SVM)

Dwi Tiyas Novitasari^{1*}; Mula Agung Barata²; Pelangi Eka Yuwita³

Teknik Informatika^{1,2}, Teknik Mesin³
Universitas Nahdlatul Ulama Sunan Giri, Bojonegoro, Indonesia^{1,2,3}
unugiri.ac.id^{1,2,3}
dwitiyasn17@gmail.com^{1*}, mula.ab26@gmail.com², pelangi.ardata@gmail.com³
(*) Corresponding Author



Ciptaan disebarluaskan di bawah Lisensi Creative Commons Atribusi-NonKomersial 4.0 Internasional.

Abstract— *The Originote Hyalucera Moisturizer skincare product has attracted public attention because it offers superior quality at an affordable price. Social media, especially Twitter, is used by consumers to express opinions regarding this product, whether positive, negative, or neutral. However, the large number of reviews with various sentiments can confuse potential consumers in assessing product quality. Therefore, this study aims to understand user perception through sentiment analysis and evaluate the effectiveness of the Support Vector Machine (SVM) algorithm in sentiment classification. A total of 1,820 tweets were collected using the crawling technique with Python. The data undergoes preprocessing, including text cleaning, tokenization, stopword removal, and stemming, reducing it to 902 tweets. Key text features are extracted using Term Frequency-Inverse Document Frequency (TF-IDF). For sentiment classification, this study used the SVM algorithm, which is known as an effective method in text processing. Model evaluation showed good results with an accuracy of 87%, precision of 89%, and recall of 87%. This study provides insight into public perception of The Originote Hyalucera Moisturizer and measures the effectiveness of SVM in social media-based sentiment analysis. The results of the study can be utilized by manufacturers for more targeted marketing strategies, product quality improvement, and more effective communication in responding to opinions on social media. In addition, this study contributes to the development of machine learning-based sentiment analysis methods in the context of skincare products.*

Keywords: *sentiment analysis, SVM, TF-IDF, twitter.*

Abstrak— Produk *skincare The Originote Hyalucera Moisturizer* menarik perhatian masyarakat karena menawarkan kualitas unggul dengan harga terjangkau. Media sosial, khususnya Twitter, digunakan oleh konsumen untuk menyampaikan opini terkait produk ini, baik positif, negatif, maupun netral. Namun, banyaknya ulasan dengan berbagai sentimen dapat membingungkan calon konsumen dalam menilai kualitas produk. Oleh karena itu, penelitian ini bertujuan untuk memahami persepsi pengguna melalui analisis sentimen serta mengevaluasi efektivitas algoritma *Support Vector Machine* (SVM) dalam klasifikasi sentimen. Sebanyak 1.820 *tweet* dikumpulkan menggunakan teknik *crawling* dengan *Python*. Data melewati tahap pra-pemrosesan, termasuk pembersihan teks, tokenisasi, penghapusan *stopword*, dan *stemming*, sehingga jumlah data yang digunakan berkurang menjadi 902 *tweet*. Fitur penting dari teks diekstrak menggunakan *Term Frequency-Inverse Document Frequency* (TF-IDF). Untuk klasifikasi sentimen, penelitian ini menggunakan algoritma SVM, yang dikenal sebagai metode efektif dalam pengolahan teks. Evaluasi model menunjukkan hasil yang baik dengan akurasi 87%, *precision* 89%, dan *recall* 87%. Penelitian ini memberikan wawasan tentang persepsi publik terhadap *The Originote Hyalucera Moisturizer* dan mengukur efektivitas SVM dalam analisis sentimen berbasis media sosial. Hasil penelitian dapat dimanfaatkan oleh produsen untuk strategi pemasaran yang lebih terarah, peningkatan kualitas produk, serta komunikasi yang lebih efektif dalam merespons opini di media sosial. Selain itu, penelitian ini berkontribusi pada pengembangan metode analisis sentimen berbasis pembelajaran mesin dalam konteks produk *skincare*.

Kata kunci: *analisis sentimen, SVM, TF-IDF, twitter.*

PENDAHULUAN

Pertumbuhan industri *skincare* di Indonesia terus meningkat seiring dengan kesadaran masyarakat terhadap pentingnya perawatan kulit. Produk lokal seperti *The Originote Hyalucera Moisturizer*, yang mengandung *Hyaluronic Acid* dan *Ceramide* untuk hidrasi maksimal, semakin diminati karena menawarkan kualitas yang kompetitif dengan harga terjangkau (Jasmarizal et al., 2024). Twitter merupakan platform sosial yang dapat dimanfaatkan siapa saja untuk bebas mengungkapkan pendapatnya (Isnain, Sakti, et al., 2021). Di sisi lain, dinamika media sosial, terutama Twitter, memainkan peran signifikan dalam membentuk persepsi masyarakat terhadap produk-produk *skincare* (Barata et al., 2024). Platform ini memungkinkan pengguna untuk berbagi ulasan, baik positif, negatif, maupun netral yang dapat memengaruhi keputusan pembelian konsumen. Namun, beragamnya opini sering kali membingungkan calon konsumen dalam menentukan pilihan (Husada & Paramita, 2021). Oleh karena itu, analisis sentimen diperlukan untuk memberikan gambaran yang lebih objektif mengenai kualitas dan efektivitas produk ini, serta memahami persepsi pengguna secara lebih mendalam.

Analisis sentimen merupakan suatu pendekatan dalam pemrosesan bahasa alami (*Natural Language Processing/NLP*) yang bertujuan untuk mengidentifikasi dan mengklasifikasikan ekspresi perasaan dalam teks ke dalam kategori positif, negatif, atau netral (Harnelia & Saputra, 2024). Menurut Liu, analisis sentimen merupakan bidang studi yang berfokus pada evaluasi opini, sentimen, sikap, dan emosi terhadap suatu entitas, seperti produk, layanan, peristiwa, organisasi, topik, serta atribut yang terkait dengannya (Isnain, Marga, et al., 2021). Sebagai bagian dari NLP, analisis sentimen berperan dalam mendeteksi subjektivitas dalam teks serta mengekstrak dan mengelompokkan opini secara sistematis (Pratiwi & Kamayani, 2024). Dalam konteks analisis sentimen, *text mining* berperan penting dalam mengidentifikasi dan mengklasifikasikan opini publik mengenai suatu produk berdasarkan data yang diperoleh dari berbagai sumber. *Text mining* merupakan proses ekstraksi informasi dari data teks yang dilakukan secara otomatis dengan tujuan menemukan pola atau wawasan baru yang sebelumnya tidak diketahui (Sitanggang et al., 2024).

Beberapa penelitian sebelumnya telah membahas analisis sentimen dalam konteks produk kecantikan, khususnya *skincare*. (Harnelia & Saputra, 2024) melakukan penelitian dengan

topik "Analisis Sentimen *Review Skincare Skintific*" menggunakan 958 data ulasan produk *skincare* dari platform media sosial. Metode *GridSearch* diterapkan untuk mengoptimalkan *Support Vector Machine* (SVM), menghasilkan tingkat akurasi sebesar 94%.

Selain itu, penelitian yang dilakukan oleh Larasati et al. (2024) dengan topik "Analisis Sentimen Produk Kecantikan Jenis *Moisturizer* di Twitter" bertujuan untuk memahami opini masyarakat terhadap produk perawatan kulit. Hasil analisis menggunakan algoritma SVM menunjukkan akurasi sebesar 98%, dengan nilai presisi dan *recall* masing-masing mencapai 98% dan 100% (Cantika Larasati et al., 2024).

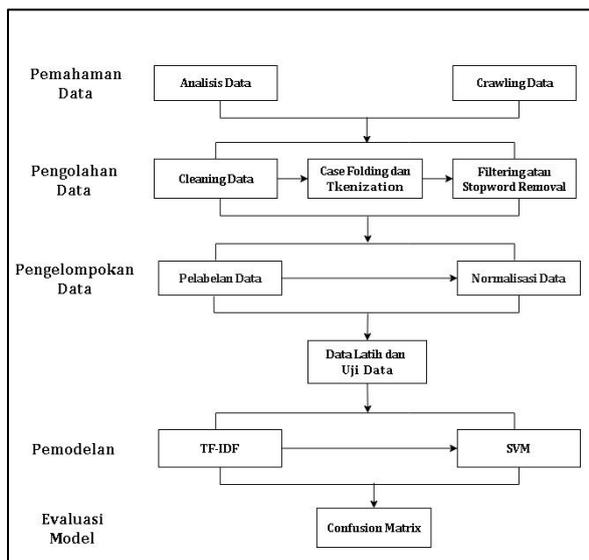
Penelitian lainnya oleh Muslimah dan Sutikno (2023) membahas sentimen terhadap merek *The Originote* melalui analisis komentar warganet. Studi ini menerapkan metode *Naïve Bayes Classifier* untuk mengelompokkan sentimen terhadap merek *skincare* tersebut. Hasil evaluasi menunjukkan tingkat akurasi sebesar 78,48%, yang mengindikasikan bahwa model yang digunakan cukup efektif dalam mengklasifikasikan opini konsumen (Muslimah & Sutikno, 2023).

Meskipun telah banyak penelitian mengenai analisis sentimen pada produk *skincare* dengan menggunakan metode *Support Vector Machine* (SVM), penerapan algoritma ini dalam mengklasifikasikan sentimen terhadap *The Originote Hyalucera Moisturizer* masih belum banyak diteliti secara mendalam. Setiap produk *skincare* memiliki karakteristik unik yang dapat memengaruhi persepsi pengguna, sehingga penting untuk meneliti sejauh mana SVM mampu mengolah dan menginterpretasikan sentimen yang muncul di platform media sosial, khususnya Twitter (Anggraini & Alita, 2024). Selain itu, faktor seperti perbedaan *dataset*, metode pra-pemrosesan, serta parameter yang digunakan dalam model SVM dapat berdampak pada hasil analisis sentimen.

Penelitian ini bertujuan untuk menganalisis sentimen review produk *skincare The Originote Hyalucera Moisturizer* dari aplikasi Twitter. Dengan mengklasifikasikan sentimen menjadi kategori positif, negatif, dan netral, penelitian ini bertujuan untuk mengevaluasi efektivitas SVM dalam menangani klasifikasi teks serta mengukur kinerjanya menggunakan metrik evaluasi seperti akurasi, presisi dan *recall*. Hasil penelitian diharapkan dapat memberikan wawasan bagi konsumen dalam mempertimbangkan keputusan pembelian serta membantu produsen memahami persepsi publik terhadap produk mereka.

BAHAN DAN METODE

Tahapan penelitian yang dilakukan dengan pendekatan analisis sentimen untuk mengkaji opini pengguna Twitter terhadap produk *skincare The Originote Hyalucera Moisturizer*. Rancangan penelitian ini mencakup pemahaman data, pengolahan data, pemodelan, dan evaluasi model. Berikut merupakan kerangka pemikiran yang dapat dilihat pada Gambar 1 dibawah ini.



Sumber: (Penelitian, 2024)
 Gambar 1. Kerangka Pemikiran

Pemahaman data

Pada tahap ini, data yang diperoleh dari *crawling* Twitter dianalisis untuk memahami konteks dan struktur dari *tweet* yang ada. Pengambilan data menggunakan *google colab*, karena platform ini menawarkan banyak fitur yang sangat berguna baik untuk pemula ataupun profesional terutama pada bidang machine learning, data science, dan pengolahan data secara umum (Laili et al., 2025). Proses ini penting untuk mengetahui jenis informasi yang terkandung dalam *tweet* dan untuk menyiapkan data agar dapat diproses lebih lanjut.

Pengolahan data

Merupakan langkah awal dalam mengubah data teks menjadi analisis sentimen melalui penggunaan alat *google colab* (Fathonah & Herliana, 2021). Data untuk penelitian ini diperoleh melalui pengumpulan sentimen warganet dari aplikasi *Twitter* menggunakan *Python*, yang dikenal sebagai bahasa pemrograman serbaguna dalam bidang *Machine Learning dan Deep Learning* (Pranata et al., 2024). Proses

crawling data berperan penting dalam ekstraksi informasi, seperti dokumen, teks, dan lainnya. Seluruh data sentimen kemudian diproses melalui tahapan pra-pemrosesan, yang meliputi penghapusan karakter khusus, *stemming*, *stopwords*, dan tokenisasi, dengan tujuan mengekstrak informasi penting dari *database*. Ada beberapa teknik yang diterapkan dalam pengolahan data, antara lain:

1. *Cleaning data* adalah prosedur mengeliminasi data teks yang tidak konsisten atau tidak relevan.
2. *Case folding* berarti mengubah bentuk kata menjadi bentuk dasarnya agar karakter memiliki keseragaman (huruf kecil). Sedangkan, *tokenizer* adalah langkah untuk membagi data teks menjadi beberapa token.
3. *Filtering StopWord* (Bahasa Inggris) adalah proses untuk menghapus kata-kata dalam bahasa Inggris.

Pelabelan data

Setelah data dibersihkan, tahap selanjutnya adalah pelabelan sentimen. Komentar pada *tweet* diklasifikasikan ke dalam tiga kategori utama: positif, negatif, dan netral. Pelabelan ini dilakukan secara semi-otomatis dengan validasi manual untuk memastikan akurasi. Proses ini sangat penting untuk mempersiapkan data agar dapat dianalisis lebih lanjut menggunakan algoritma *machine learning*.

Pembobotan TF-IDF

Setelah data dilabeli, tahap berikutnya adalah pemberian bobot pada setiap kata dalam data menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF). Metode ini digunakan untuk menghitung relevansi atau kepentingan setiap kata dalam analisis, dengan tujuan meningkatkan akurasi model dalam mengidentifikasi sentimen (Ashari et al., 2023).

Pemodelan Algoritma svm

Pada tahap ini, data yang telah diberi bobot TF-IDF digunakan untuk melatih model menggunakan algoritma *Support Vector Machine* (SVM). *Support Vector Machine* adalah teknik yang sering dipakai untuk mengkategorikan data teks dengan tingkat ketepatan yang lebih baik (Ardiansyah et al., 2023). SVM termasuk dalam kategori teknik pembelajaran terawasi yang diterapkan saat melakukan klasifikasi. Dalam kerangka klasifikasi, SVM memiliki prinsip-prinsip matematis yang lebih solid dan terperinci. SVM digunakan untuk menemukan *hyperplane* terbaik dengan meningkatkan jarak antara kelas-kelas.

Evaluasi model

Proses evaluasi model dilakukan menggunakan berbagai metrik untuk mengukur performa model dalam mengklasifikasikan sentimen dengan akurat. Metrik yang digunakan meliputi akurasi, presisi, dan *recall*. Evaluasi ini penting untuk mengetahui sejauh mana model dapat mengklasifikasikan komentar-komentar terkait produk *The Originote Hyalucera Moisturizer* dengan tepat.

HASIL DAN PEMBAHASAN

Pemahaman Data

Langkah pertama dalam mengembangkan solusi untuk analisis sentimen mengenai *The Originote Hyalucera Moisturizer* dengan memanfaatkan data dari Twitter adalah melakukan pemahaman yang mendalam tentang objek penelitian. Data yang diambil berupa cuitan-cuitan yang berhubungan dengan produk ini, yang mungkin mengandung sentimen, pendapat, atau pandangan dari pengguna. Tujuan dari analisis sentimen ini adalah untuk menentukan apakah pengguna Twitter memberikan sentimen yang positif, negatif, atau netral terhadap produk skincare ini. Informasi yang diperoleh akan sangat berguna bagi perusahaan dalam memahami perspektif konsumen, menilai elemen yang mendapat pujian atau kritik, serta mendukung rencana pemasaran dan branding.

Pada tahap ini, pengumpulan data dilakukan melalui proses *crawling* data menggunakan Twitter API, yang memungkinkan pengambilan cuitan dalam jumlah besar berdasarkan kata kunci atau tagar tertentu, seperti *The Originote Hyalucera Moisturizer* atau *#OriginoteSkincare*. Proses ini mencakup langkah-langkah teknis, mulai dari autentikasi API hingga penerapan filter untuk memastikan relevansi data. Data yang diambil kemudian diformat menjadi *file* terstruktur seperti CSV. Hasil *crawling* ini menjadi fondasi bagi analisis lebih lanjut untuk mengidentifikasi pola sentimen dan tren persepsi publik secara keseluruhan. Berikut adalah tahapan pengambilan dan pengolahan data yang dilakukan oleh peneliti:

1. Kata kunci yang digunakan dalam proses *crawling* adalah *The Originote Hyalucera Moisturizer*. Dari hasil *crawling*, diperoleh total 1.820 data, yang kemudian melalui proses pembersihan, menghasilkan 920 data yang siap digunakan.
2. Data sentimen yang dianalisis berasal dari unggahan pengguna media sosial Twitter berbahasa Indonesia, yang dikumpulkan melalui proses *crawling* menggunakan bahasa

pemrograman Python pada Google Colab. Setelah data berhasil diambil, setiap entri diberi label positif, negatif, atau netral, lalu dilakukan proses normalisasi sesuai dengan standar Kamus Besar Bahasa Indonesia (KBBI).

Pengolahan Data

Pengolahan data dalam penelitian ini merupakan langkah penting untuk memastikan data siap digunakan dalam analisis sentimen. Proses ini dimulai dengan *cleansing* data, yaitu membersihkan teks dari elemen-elemen yang tidak relevan seperti simbol, emotikon, URL, tanda baca berlebihan, dan kata-kata kasar. Langkah ini bertujuan untuk memastikan teks hanya memuat informasi yang penting dan relevan untuk analisis. Selanjutnya, data diproses melalui *folding* dan *tokenization*, dimana seluruh teks diubah menjadi huruf kecil (*lowercase*) untuk menghindari duplikasi kata akibat perbedaan kapitalisasi, kemudian teks dipecah menjadi token atau kata-kata individu. Proses ini membantu model analisis memahami setiap komponen teks dengan lebih baik. Setelah itu, dilakukan *filtering* dan *stopword removal* untuk menghapus kata-kata yang tidak memiliki nilai analitis, seperti kata-kata umum atau *stopwords* (contoh: "yang", "dan", "gw" "dari"). Dengan menghilangkan kata-kata tersebut, data menjadi lebih fokus pada kata-kata yang bermakna dalam menentukan sentimen.

1. Data Cleaning

Proses ini berfungsi membersihkan, *replace* data atau karakter yang tidak baku yang dapat mengganggu pengolahan data yang dilakukan pada Google Colab, seperti terlihat pada Tabel 1.

Tabel 1. *Data Cleaning*

No	<i>full_text</i>	<i>Cleaning</i>
1	The Originote Hyalucera Moisturizer Gel - Pele...	The Originote Hyalucera Moisturizer Gel Pelem...
2	WTS moisturizer ceratinol & hyalucera The ...	WTS moisturizer ceratinol amp hyalucera The Or...
3	Promo Payday Sale The Originote Hyalucera Mois...	Promo Payday Sale The Originote Hyalucera Mois...
4	The Originote Hyalucera Moisturizer Gel 50 ml ... Cek The Originote	The Originote Hyalucera Moisturizer Gel ml P... Cek The Originote
5	Hyalucera Moisturizer Ge	Hyalucera Moisturizer Gel P...

Sumber: (Penelitian, 2024)

2. Folding dan Tokenization

Proses ini bertujuan untuk membagi teks atau dokumen menjadi unit-unit kecil yang disebut

"token". Langkah ini dilakukan agar analisis teks menjadi lebih mudah dengan mengonversinya ke dalam format yang lebih terstruktur dan dapat diolah secara efisien oleh komputer. Hasil dari tahap *Tokenize* pada Tabel 2.

Tabel 2. *Folding dan Tokenization*

No	full_text	Cleaning	Tokenization and Case Folding
1	The Originote Hyalucera Moisturizer Gel - Pele...	The Originote Hyalucera Moisturizer Gel Pelem...	[the, originote, hyalucera, moisturizer, gel, ...
2	WTS moisturizer ceratinol & hyalucera The ...	WTS moisturizer ceratinol amp hyalucera The Or...	[wts, moisturizer, ceratinol, amp, hyalucera, ...
3	Promo Payday Sale The Originote Hyalucera Mois...	Promo Payday Sale The Originote Hyalucera Mois...	[promo, payday, sale, the, originote, hyalucera...
4	The Originote Hyalucera Moisturizer Gel 50 ml ...	The Originote Hyalucera Moisturizer Gel ml P...	[the, originote, hyalucera, moisturizer, gel, ...
5	Cek The Originote Hyalucera Moisturizer Ge -	Cek The Originote Hyalucera Moisturizer Gel P...	[cek, the, originote, hyalucera, moisturizer, ...

Sumber: (Penelitian, 2024)

3. *Filtering atau Stopword Removal*

Tahap akhir dalam proses pra-pemrosesan adalah penyaringan *stopwords*. Pada tahap ini, kata-kata yang memiliki sedikit makna atau tidak berkontribusi secara signifikan terhadap analisis akan dihapus dari teks. Hasil dari penyaringan *stopwords* menggunakan Google Colab ditampilkan dalam Tabel 3.

Tabel 3. *Filtering atau Stopword Removal*

No	full_text	cleaning	Tokenization and Case Folding	Filtering/stopword removal
1	The Originote Hyalucera Moisturizer Gel - Pele...	The Originote Hyalucera Moisturizer Gel Pelem...	[the, originote, hyalucera, moisturizer, gel, ...	[the, originote, hyalucera, moisturizer, gel, ...
2	WTS moisturizer ceratinol & hyalucera The ...	WTS moisturizer ceratinol amp hyalucera The Or...	[wts, moisturizer, ceratinol, amp, hyalucera, ...	[wts, moisturizer, ceratinol, amp, hyalucera, ...
3	Promo Payday Sale The	Promo Payday Sale The Originote	[promo, payday, sale, the, originote,	[promo, payday, sale, the,

No	full_text	cleaning	Tokenization and Case Folding	Filtering/stopword removal
	Originote Hyalucera Mois...	Hyalucera Mois...	hyalucera...	originote, hyalucera...
4	The Originote Hyalucera a Moisturizer Gel 50 ml ...	The Originote Hyalucera Moisturizer Gel ml P...	[the, originote, hyalucera, moisturizer, gel, ...	[the, originote, hyalucera, moisturizer, gel, ...
5	Cek The Originote Hyalucera a Moisturizer Ge -	Cek The Originote Hyalucera Moisturizer Gel P...	[cek, the, originote, hyalucera, moisturizer, ..	[cek, the, originote, hyalucera, moisturizer, ...

Sumber: (Penelitian, 2024)

Pada tahap ini, kata-kata yang umumnya memiliki nilai informasi rendah dalam sebuah teks akan dihapus.

Pemodelan

Algoritma yang akan digunakan pada penelitian ditentukan pada tahap ini. Algoritma SVM. Penelitian ini menggunakan pendekatan TF-IDF untuk mempresentasikan data teks ke dalam bentuk vektor numerik yang dapat diproses oleh algoritma. Ekstraksi fitur dengan TF-IDF digunakan untuk mentransformasikan dokumen ulasan ke pada struktur yang bisa dikenakan untuk prosedur klasifikasi dengan SVM, yaitu dalam bentuk vektor. Hal ini dilakukan pembobotan kata dalam setiap fitur yang akan dihitung kemunculan dalam dokumen disebut TF. Lalu, akan dihitung nilai IDF, dan yang terakhir akan dihitung TF-IDF. Model rancangan penelitian pada Google Colab ditampilkan pada Gambar 2.

```
#@title Pembobotan TF-IDF
from sklearn.feature_extraction.text import TfidfVectorizer

# inisialisasi TF-IDF Vectorizer
tfidf_vectorizer = TfidfVectorizer()

# Transformasi data teks menjadi vektor TF-IDF
X_train_tfidf = tfidf_vectorizer.fit_transform(X_train)
X_test_tfidf = tfidf_vectorizer.transform(X_test)

# Cetak nama fitur dari TF-IDF Vectorizer
Feature_names = tfidf_vectorizer.get_feature_names_out()

# Cetak beberapa contoh hasil transformasi dari data pelatihan
print("Hasil transformasi dari data pelatihan:")
print(X_train_tfidf)

# Cetak beberapa contoh fitur dari vektor TF-IDF
print("\nfitur dari vektor TF-IDF:")
print(Feature_names)
```

Sumber: (Hasil Penelitian, 2024)

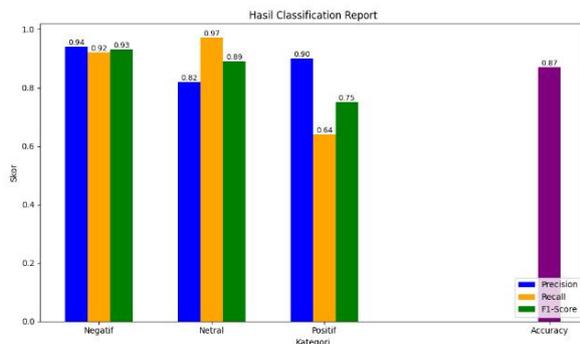
Gambar 2. Rancangan Model Peneliti

Pelatihan SVM merupakan tahapan pembentuk klasifikasi. Mengubah semua fitur data latih berlabel positif di beri kelas 1, data label negatif diberi kelas -1, dan data label netral diberi kelas 0. Data latih yang sudah diubah menjadi format *support vector*. Kemudian, dilakukan perhitungan kernelisasi menggunakan kernel linier. Berdasarkan hasil perhitungan fungsi keputusan klasifikasi Data uji ini diklasifikasikan ke dalam kelas netral (0) karena nilai fungsi keputusan untuk kelas netral memiliki skor tertinggi.

Evaluasi Model

Evaluasi model dalam penelitian ini merupakan tahap di mana kinerja sistem analisis sentimen diukur untuk menilai sejauh mana model mampu memprediksi atau mengklasifikasikan sentimen dalam teks secara akurat. Untuk menilai performa model, digunakan beberapa metrik evaluasi, yaitu akurasi, *precision* dan *recall*.

Pengujian dilakukan menggunakan Google Colab, yang dirancang untuk mengevaluasi efektivitas model dan alur kerja (*workflow*) yang telah dikembangkan. Proses ini bertujuan untuk menguji seberapa baik model yang telah disusun dalam memprediksi atau menganalisis sentimen berdasarkan data yang digunakan. Gambar 3 menampilkan hasil evaluasi model yang divisualisasikan dalam laporan klasifikasi (*Classification Report*).



Sumber: (Hasil Penelitian, 2024)
 Gambar 3. Hasil Evaluasi Model

Dari serangkaian pengujian yang dilakukan, diperoleh hasil evaluasi berupa nilai akurasi, *precision*, *recall*, dan *confusion matrix*.

1. Precision, Recall dan Akurasi

Berdasarkan hasil yang ditampilkan dalam Tabel 4, model yang diuji menunjukkan tingkat akurasi yang baik, yaitu sebesar 87%, yang mengindikasikan bahwa model memiliki performa yang cukup optimal dalam mengklasifikasikan sentimen dengan tingkat kesalahan yang rendah.

Tabel 4. Hasil *precision*, *Recall* dan Akurasi

No	Keterangan	Hasil
1	<i>Precision</i>	0.88 %
2	<i>Recall</i>	0.87 %
3	<i>Akurasi</i>	0.87 %

Sumber: (Hasil Penelitian, 2024)

2. Confusion Matrix

Hasil dari *confusion matrix* pengujian pada penelitian ini ditampilkan pada Gambar 4.

Aktual \ Prediksi	Negatif	Netral	Positif
Negatif	84	6	1
Netral	0	130	4
Positif	5	22	47

Sumber: (Hasil Penelitian, 2024)

Gambar 4. Confusion Matrix

Dari gambar *confusion matrix* pada Gambar 4, model menghasilkan akurasi sebesar 87%, menunjukkan bahwa dari total data uji, 87% klasifikasi sentimen dilakukan dengan benar. Dari tabel *confusion matrix*, terlihat bahwa kelas netral memiliki performa terbaik dibanding kelas lainnya, dengan jumlah prediksi benar.

Dibandingkan dengan penelitian sebelumnya (Harnelia & Saputra, 2024), ditemukan bahwa akurasi model pada penelitian menunjukkan tingkat akurasi sebesar 94%, dengan distribusi performa yang berbeda setiap kelas. Faktor yang dapat mempengaruhi perbedaan ini antara lain teknik *preprocessing*, pelabelan, metode klasifikasi yang diterapkan, serta jumlah dan distribusi data dalam setiap kelas. Perbedaan dalam teknik ekstraksi fitur, seperti penggunaan TF-IDF juga dapat mempengaruhi.

Kelebihan model dalam penelitian ini adalah kemampuan yang sangat baik mengidentifikasi komentar netral, yang terlihat dari *recall* tinggi 97%.

Hal ini menunjukkan bahwa komentar dengan nada netral memiliki pola bahasa yang lebih mudah dikenali oleh model. Namun, kelemahan utama model ini adalah akurasi yang lebih rendah dalam klasifikasi sentimen positif, yang kemungkinan disebabkan oleh kemiripan pola

bahasa antara sentimen positif dan netral. Selain itu, jumlah data positif yang lebih sedikit dibandingkan netral dan negatif dapat menjadi penyebab model mengalami kesulitan dalam membedakan komentar positif secara lebih akurat.

KESIMPULAN

Penerapan algoritma SVM dengan pendekatan TF-IDF terbukti efektif dalam mengklasifikasikan sentimen pengguna Twitter terhadap produk *skincare The Originote Hyaluronic Moisturizer*. Model yang dibangun berhasil mengidentifikasi sentimen positif, negatif, dan netral dengan tingkat akurasi sebesar 87%, *precision* 89%, dan *recall* 87%. Hasil ini mengindikasikan bahwa metode yang digunakan memiliki kemampuan yang andal dalam memahami persepsi publik berdasarkan analisis teks dari media sosial. Penelitian ini memberikan kontribusi signifikan bagi produsen, khususnya dalam menyusun strategi pemasaran yang lebih efektif, meningkatkan kualitas produk sesuai dengan masukan konsumen, dan merancang komunikasi yang mampu menghadapi opini negatif dengan lebih baik. Selain itu, penelitian ini memperkuat relevansi analisis sentimen berbasis media sosial sebagai pendekatan modern yang dapat digunakan dalam memahami tren opini masyarakat terhadap produk perawatan kulit.

REFERENSI

- Anggraini, J., & Alita, D. (2024). Implementasi Metode SVM Pada Sentimen Analisis Terhadap Pemilihan Presiden (Pilpres) 2024 Di Twitter. *Jurnal Informatika: Jurnal Pengembangan IT*, 9(2), 102–111. <https://doi.org/10.30591/jpit.v9i2.6560>
- Ardiansyah, D., Saepudin, A., Aryanti, R., Fitriani, E., & Royadi. (2023). Analisis Sentimen Review Pada Aplikasi Media Sosial Tiktok Menggunakan Algoritma K-Nn Dan Svm Berbasis Pso. *Jurnal Informatika Kaputama (JIK)*, 7(2), 233–241. <https://doi.org/10.59697/jik.v7i2.148>
- Ashari, S. A., Saputra, M. W. A., Larosa, E., & Rijal, B. S. (2023). Analisis Sentimen pada Aplikasi Translate Google Menggunakan Metode SVM (Studi Kasus: Komentar Pada Playstore). *Jurnal Teknik*, 21(2), 168–182. <https://doi.org/10.37031/jt.v21i2.412>
- Barata, M., Ayuni, I. S., Kartini, A. Y., & Alawi, Z. (2024). Algoritma K-Means dalam Clustering Produk Skincare untuk Menentukan Strategi Pemasaran. *Jurnal Informatika Polinema*, 10(3), 421–428.

- <https://doi.org/10.33795/jip.v10i3.5167>
- Cantika Larasati, R., Dewi, C., & Juli, C. H. (2024). Analisis sentimen produk kecantikan jenis moisturizer di twitter menggunakan algoritma super vector machine. *Teknikom*, 7(1), 124–134. <https://doi.org/10.37600/tekinkom.v7i1.1243>
- Fathonah, F., & Herliana, A. (2021). Penerapan Text Mining Analisis Sentimen Mengenai Vaksin Covid - 19 Menggunakan Metode Naive Bayes. *Jurnal Sains Dan Informatika*, 7(2), 155–164. <https://doi.org/10.34128/jsi.v7i2.331>
- Harnelia, H., & Saputra, R. A. (2024). Analisis Sentimen Review Skincare Skintific Dengan Algoritma Support Vector Machine (Svm). *Jurnal Informatika Dan Teknik Elektro Terapan*, 12(2). <https://doi.org/10.23960/jitet.v12i2.4095>
- Husada, H. C., & Paramita, A. S. (2021). Analisis Sentimen Pada Maskapai Penerbangan di Platform Twitter Menggunakan Algoritma Support Vector Machine (SVM). *Teknika*, 10(1), 18–26. <https://doi.org/10.34148/teknika.v10i1.311>
- Isnain, A. R., Marga, N. S., & Alita, D. (2021). Sentiment Analysis Of Government Policy On Corona Case Using Naive Bayes Algorithm. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, 15(1), 55. <https://doi.org/10.22146/ijccs.60718>
- Jasmarizal, Junadhi, Rahmadden, & M. Khairul Anam. (2024). Penerapan Metode Support Vector Machine Untuk Analisis Sentimen Terhadap Produk Skincare. *Indonesian Journal of Computer Science*, 13(1), 1438–1450. <https://doi.org/10.33022/ijcs.v13i1.3654>
- Laili, E., Alawi, Z., Rohmah, R., & Barata, M. (2025). KOMPARASI ALGORITMA DECISION TREE DAN SUPPORT VECTOR MACHINE (SVM) DALAM KLASIFIKASI SERANGAN JANTUNG. *Jurnal Sistem Informatika Dan Informatika (Simika)*, 8(1), 67–76. <https://doi.org/10.47080/simika.v8i1.3683>
- Muslimah, N., & Sutikno, A. (2023). Analisis Sentimen Komentar Netizen Pada Brand Skincare The Originote Menggunakan Metode Naive Bayes. *Proceedings of the National Conference on Electrical Engineering, Informatics, Industrial Technology, and Creative Media*, 3(1), 826–835. Retrieved from <https://conferences.ittelkom-pwt.ac.id/index.php/centive/article/view/131>
- Pranata, R. A., Rudiman, R., & Verdikha, N. A. (2024). METODE PEMBOBOTAN TF-IDF

UNTUK KLASIFIKASI TEKS QUICK COUNT PEMILIHAN WAKIL PRESIDEN INDONESIA 2024 PADA X TWITTER DENGAN METODE SVM. *Jurnal Teknologi Informasi : Jurnal Keilmuan dan Aplikasi Bidang Teknik Informatika*, 18(2), 126-138. <https://doi.org/10.47111/jti.v18i2.14934>

Pratiwi, A., & Kamayani, M. (2024). Perbandingan Pelabelan Data dalam Analisis Sentimen Kurikulum Proyek di platform TikTok: Pendekatan Naïve Bayes. *Jurnal Eksplora Informatika*, 14(1), 96-107. <https://doi.org/10.30864/eksplora.v14i1.1093>

Sitanggang, A., Umaidah, Y., Umaidah, Y., Adam, R. I., & Adam, R. I. (2024). Analisis Sentimen Masyarakat Terhadap Program Makan Siang Gratis Pada Media Sosial X Menggunakan Algoritma Naïve Bayes. *Jurnal Informatika Dan Teknik Elektro Terapan*, 12(3). <https://doi.org/10.23960/jitet.v12i3.4902>