



DOI: <https://doi.org/10.38035/jim.v4i3>
<https://creativecommons.org/licenses/by/4.0/>

Memanfaatkan Analisis Sentimen Twitter untuk Mengelola Reputasi Merek: Studi Kasus *Skincare* Merek X Menggunakan *Support Vector Machine*

Muhammad Wildan Agba¹, Puspa Eosina², Dewi Primasari³

¹Program Studi Teknik Informatika, Fakultas Teknik dan Sains Universitas Ibn Khaldun, Bogor, Indonesia, wildan.agba@gmail.com

²Program Studi Teknik Informatika, Fakultas Teknik dan Sains Universitas Ibn Khaldun, Bogor, Indonesia, puspa.eosina@ft.uika-bogor.ac.id

³Program Studi Teknik Informatika, Fakultas Teknik dan Sains Universitas Ibn Khaldun, Bogor, Indonesia, dewi.primasari@uika-bogor.ac.id

Corresponding Author: wildan.agba@gmail.com¹

Abstract: *This research evaluates the reputation of Merek X by analyzing 2,646 tweets from October 2023 to March 2024 using natural language processing (NLP) and Support Vector Machine (SVM). This study aims to assess consumer sentiment towards Merek X, identify the main factors influencing sentiment, and measure the data distribution and effectiveness of the SVM algorithm. The results showed that 69% of the sentiment was positive, 18.2% negative, and 12.8% neutral, with the SVM model achieving 83% accuracy. Factors such as product fit, price, quality, and innovation were found to significantly influence sentiment. These findings suggest that Merek X can improve customer satisfaction and refine its marketing strategy to strengthen its market position.*

Keyword: *Brand Reputation; Natural Language Processing; Sentiment Analysis; Support Vector Machine; Skincare.*

Abstrak: Penelitian ini mengevaluasi reputasi Merek X dengan menganalisis 2.646 tweet dari Oktober 2023 hingga Maret 2024 menggunakan natural language processing (NLP) dan Support Vector Machine (SVM). Penelitian ini bertujuan untuk menilai sentimen konsumen terhadap Merek X, mengidentifikasi faktor-faktor utama yang mempengaruhi sentimen, serta mengukur distribusi data, dan keefektifan algoritma svm. Hasil penelitian menunjukkan bahwa 69% dari sentimen adalah positif, 18,2% negatif, dan 12,8% netral, dengan model SVM mencapai akurasi 83%. Faktor-faktor seperti kecocokan produk, harga, kualitas, dan inovasi ditemukan secara signifikan mempengaruhi sentimen. Temuan ini menunjukkan bahwa Merek X dapat meningkatkan kepuasan pelanggan dan menyempurnakan strategi pemasarannya untuk memperkuat posisi pasarnya.

Kata Kunci: Reputasi Merek; Pemrosesan Bahasa Alami; Analisis Sentimen; *Support Vector Machine*; *Skincare*.

PENDAHULUAN

Dalam beberapa tahun terakhir, industri kosmetik di Indonesia telah berkembang pesat, terutama di segmen perawatan kulit. Hingga Oktober 2023, jumlah perusahaan kosmetik meningkat menjadi 1.010, dari 200 perusahaan di tahun 2020 (Waluyo et al., 2023). Perubahan kebiasaan konsumen akibat pandemi COVID-19 juga berkontribusi terhadap peningkatan penggunaan produk perawatan kulit di rumah (Mutia, 2022; statista, 2023). Dalam persaingan industri yang kompetitif ini, reputasi merek menjadi faktor utama dalam keputusan pembelian (Lin et al., 2023; Mamuaya, 2024; Mohammad et al., 2015; Traore, 2024; ZAP Clinic and MarkPlus, 2024). Reputasi yang kuat menumbuhkan kepercayaan konsumen dengan menekankan kejujuran dan kualitas (Fatmawati and Fauzan, 2021; Mohammad et al., 2015; Stravinskiene et al., 2021). Sentimen di media sosial, performa produk, dan kualitas merupakan elemen penting yang membentuk reputasi tersebut (Mostaghel et al., 2019; Olaleye et al., 2018; Purba et al., 2022; Suryavanshi and Gangele, 2024).

Merek X adalah pendatang baru di pasar skincare Indonesia, yang mulai beroperasi pada tahun 2022. Perusahaan ini menawarkan rangkaian produk skincare dengan harga terjangkau yang dirancang untuk mengatasi berbagai masalah kulit. Inovasi yang dilakukan terus-menerus dalam produk facial wash, toner, serum, dan moisturizer adalah hal yang membedakan Merek X dalam industri yang sangat kompetitif ini. Seiring dengan perkembangan dan inovasi yang terjadi, penting untuk mengevaluasi bagaimana produk tersebut dipersepsikan oleh pelanggan dan bagaimana reputasi merek tersebut berkembang di pasar (Kayakuş et al., 2024). Dalam konteks ini, melakukan analisis sentimen di Twitter sangat menguntungkan. Data Twitter yang sangat luas dan arus informasi yang cepat memungkinkan untuk mendapatkan wawasan cepat tentang pendapat konsumen mengenai suatu produk. Karena pengguna sering membagikan pengalaman mereka secara terbuka, menjadikan Twitter sebagai sumber daya yang sangat berharga untuk menilai sentimen publik terhadap merek dan produk (Khuspe et al., 2024; V. Singh et al., 2024).

Penelitian ini bertujuan untuk menganalisis sentimen konsumen terhadap skincare merek X dengan menilai representasi sentimen, mengidentifikasi faktor-faktor kunci yang mempengaruhi sentimen, mengukur distribusi data, dan mengevaluasi keefektifan metode Support Vector Machine (SVM) dalam analisis sentimen. Menganalisis sentimen konsumen terhadap merek X dengan mengidentifikasi faktor-faktor kunci yang mempengaruhi sentimen, memberikan wawasan yang dapat ditindaklanjuti untuk meningkatkan keterlibatan pelanggan. Tujuan ini sangat penting untuk memahami reputasi merek dan pengembangan produk serta memandu keputusan pemasaran strategis.

Penelitian ini menggunakan Support Vector Machine (SVM) untuk mengklasifikasikan sentimen di Twitter sebagai positif, negatif, atau netral. SVM dipilih karena telah terbukti mampu menangani data dalam jumlah besar dengan tingkat akurasi yang tinggi, berkisar antara 82% hingga 98% (Fadlil et al., 2024; Ganga et al., 2022; Imanuddin et al., 2023; Jannah and Kusnawi, 2024; Makhtum and Muhajir, 2023; Qaiser et al., 2021; Sudhakaran and Jaiganesh, 2020). Metode ini sangat efektif dalam mengelola data yang kompleks dan memberikan hasil yang dapat diandalkan, sehingga menjadi pilihan yang sangat baik untuk mempelajari reputasi merek dalam konteks yang dinamis.

Selain memberikan gambaran umum yang komprehensif tentang opini konsumen, teknik ini juga menawarkan wawasan strategis yang berharga untuk pengembangan produk dan inisiatif pemasaran Merek X di masa depan (Dubey and Jain, 2019; Fadlil et al., 2024; Rimpay et al., 2024; N. Singh and Jaiswal, 2023; Tellawar et al., 2023; Vaibhavi et al., 2024; Xi, 2024).

METODE

Pengumpulan Data

Data dikumpulkan di Twitter dari Oktober 2023 hingga Maret 2024 menggunakan pustaka Tweet-Harvest dan kata kunci “Merek X”. Jumlah total tweet yang diambil adalah 2.646, yang mencakup teks, ID diskusi, tanggal, jumlah favorit, retweet, informasi pengguna, dan lain sebagainya.

Pra-pemrosesan Data.

Pra-pemrosesan data adalah langkah penting dalam analisis sentimen, untuk memastikan kualitas dan relevansi data yang dianalisis. Metode pra-pemrosesan studi ini dirinci di bawah ini:

1. Pembersihan Data

Penanganan Data yang Tidak Lengkap: Tweet yang tidak memiliki teks akan dihapus dari kumpulan data. Penghapusan Karakter yang Tidak Relevan: Ekspresi reguler digunakan untuk menghapus karakter khusus, emotikon, emoji, URL, hashtag, dan Sebutan yang tidak relevan dengan analisis sentimen. Penghapusan Duplikasi: Tweet yang terduplikasi dihapus untuk menjamin bahwa setiap informasi adalah unik dan tidak memengaruhi analisis. Filter Bahasa: Hanya tweet dalam bahasa Indonesia yang dianalisis; tweet dalam bahasa lain diabaikan. Pelipatan Huruf (Case Folding): Semua teks diubah menjadi huruf kecil. contoh prosesnya dapat dilihat pada **Tabel 1**.

Tabel 1. Contoh Pembersihan Data

Tweet Asli	Tweet bersih
@merekX produk yang bagus tapi pengirimannya lama sekali!   #kecewa https://t.co/link1234	produk bagus tapi pengirimannya lama sekali

2. Normalisasi Teks

Kata-kata slang bahasa Indonesia yang umum dikonversi ke dalam bahasa standar dengan menggunakan kamus (KBBI) yang telah ditetapkan. contoh prosesnya dapat dilihat pada **Tabel 2**.

Tabel 2. Contoh Normalisasi Teks

Tweet Asli	Tweet Normalisasi
produk bagus tapi pengirimannya lemot	produk bagus tapi pengirimannya lambat

3. Stop word removal dan stemming

Stop word Removal: Daftar stop word bahasa Indonesia digunakan untuk menghilangkan istilah umum yang tidak memiliki arti penting. Kata-kata umum yang tidak memberikan konteks sentimen yang berharga (misalnya, “dan”, “di”, “yang”) akan dihapus. Pembatasan kata (stemming): Kata-kata direduksi menjadi bentuk dasarnya dengan menggunakan algoritma stemming berdasarkan Kamus Besar Bahasa Indonesia (KBBI). untuk menyederhanakan teks, contoh prosesnya dapat dilihat pada **Tabel 3**.

Tabel 3. Contoh Stemming

Tweet Asli	Tweet Stemming
produk bagus tapi pengirimannya lambat	produk bagus tapi kirim lambat

Pelabelan Data

Data diberi label menggunakan Kamus Leksikon Sentimen Indonesia. Kamus ini memberikan skor sentimen pada setiap tweet, yang menunjukkan apakah tweet tersebut positif, netral, atau negatif. Proses pelabelan terdiri dari menemukan kata kunci dan frasa dengan makna sentimen tertentu dan menghitung skor sentimen keseluruhan untuk setiap tweet. Beberapa penelitian telah menilai dan mengonfirmasi bahwa leksikon merupakan alat

analisis sentimen yang efektif. dimana memberikan skor sentimen, jika skornya (>0) , sentimen ditandai sebagai positif; jika (<0), ditandai sebagai negatif; jika nol, netral.

Rekayasa Fitur

Fitur diproses dengan menerapkan ekstraksi fitur dan seleksi fitur. Ekstraksi Fitur: Menggunakan metode Term Frequency-Inverse Document Frequency (TF-IDF) dan n-gram untuk menangkap konteks kata dalam tweet; Seleksi Fitur: Menggunakan uji chi-square untuk menentukan fitur yang paling relevan dengan kategori sentimen.

1. Ekstraksi Fitur: Contoh TF-IDF dan N-Gram (1,1)

Menghitung term frequency (TF), yaitu berapa kali sebuah kata muncul dalam sebuah dokumen dibagi dengan jumlah total kata dalam dokumen tersebut. contoh prosesnya dapat dilihat pada **Tabel 4**.

Contoh Dokumen:

Dokumen 1 (D1): “scarlet mahal tim sukses originote”

Dokumen 2 (D2): “produk bagus tapi pengirimannya lama”

Dokumen 3 (D3): “micellar water originote”

Tabel 4. Contoh penghitungan term frequency (TF)

Term	D1	D2	D3
scarlet	1/5	0	0
mahal	1/5	0	0
tim	1/5	0	0
sukses	1/5	0	0
originote	1/5	0	1/3
produk	0	1/5	0
bagus	0	1/5	0
tapi	0	1/5	0
pengirimannya	0	1/5	0
lama	0	1/5	0
micellar	0	0	1/3
water	0	0	1/3

IDF mengukur seberapa penting sebuah kata di seluruh dokumen. Ini dihitung sebagai:

$$IDF(t) = \log \left(\frac{N}{df(t)} \right) \quad (1)$$

Dimana:

N = adalah jumlah total dokumen.

$df(t)$ = adalah jumlah dokumen yang mengandung term t

Untuk setiap term, hitung frekuensi dokumen $df(t)$ dan kemudian IDF contoh prosesnya dapat dilihat pada **Tabel 5**.

Tabel 5. Contoh penghitungan inverse document frequency (IDF)

Term	$df(t)$	IDF
scarlet	1	$\log (3/1) = 0,477$
mahal	1	$\log (3/1) = 0,477$
tim	1	$\log (3/1) = 0,477$
sukses	1	$\log (3/1) = 0,477$
originote	2	$\log (3/2) = 0,176$
produk	1	$\log (3/1) = 0,477$
bagus	1	$\log (3/1) = 0,477$
tapi	1	$\log (3/1) = 0,477$

pengirimannya	1	$\log(3/1) = 0,477$
lama	1	$\log(3/1) = 0,477$
micellar	1	$\log(3/1) = 0,477$
water	1	$\log(3/1) = 0,477$

Terakhir, hitung TF-IDF untuk setiap term pada setiap dokumen dengan mengalikan term frequency (TF) dengan inverse document frequency (IDF), contoh prosesnya dapat dilihat pada **Tabel 6**.

Tabel 6. Contoh menghitung TF-IDF

Term	D1(TF-IDF)	D2(TF-IDF)	D3(TF-IDF)
scarlet	$1/5 \times 0,477 = 0,0954$	0	0
mahal	$1/5 \times 0,477 = 0,0954$	0	0
Tim	$1/5 \times 0,477 = 0,0954$	0	0
sukses	$1/5 \times 0,477 = 0,0954$	0	0
originote	$1/5 \times 0,176 = 0,0352$	0	$1/3 \times 0,176 = 0,0587$
produk	0	$1/5 \times 0,477 = 0,0954$	0
bagus	0	$1/5 \times 0,477 = 0,0954$	0
tapi	0	$1/5 \times 0,477 = 0,0954$	0
pengirimannya	0	$1/5 \times 0,477 = 0,0954$	0
lama	0	$1/5 \times 0,477 = 0,0954$	0
micellar	0	0	$1/3 \times 0,477 = 0,1590$
water	0	0	$1/3 \times 0,477 = 0,1590$

2. Seleksi Fitur: Contoh Chi-Square

Pertama-tama Hitung Frekuensi Observasi, buatlah tabel kontingensi yang menunjukkan kehadiran (1) atau ketidakhadiran (0) dari setiap istilah di semua dokumen, contoh prosesnya dapat dilihat pada Tabel 7.

Tabel 7. Contoh menghitung frekuensi observasi

Term	D1	D2	D3	Total
scarlet	1	0	0	1
mahal	1	0	0	1
Tim	1	0	0	1
sukses	1	0	0	1
originote	1	0	1	2
produk	0	1	0	1
bagus	0	1	0	1

tapi	0	1	0	1
pengirimannya	0	1	0	1
lama	0	1	0	1
micellar	0	1	0	1
water	0	0	1	1

Menghitung expected frequency. expected frequency dihitung dengan menggunakan:

$$E_{ij} = \left(\frac{R_i \times C_j}{N} \right) \quad (2)$$

R_i adalah jumlah total dokumen dalam baris ke- i

C_j adalah jumlah total kemunculan term ke- j .

N adalah jumlah total kemunculan term pada semua dokumen.

Sebagai contoh, expected frequency untuk kata scarlet

$$E_{\text{scarlet}, D1} = \left(\frac{1 \times 5}{13} \right) = 0.3846$$

$$E_{\text{scarlet}, D2} = \left(\frac{0 \times 6}{13} \right) = 0$$

$$E_{\text{scarlet}, D3} = \left(\frac{0 \times 2}{13} \right) = 0$$

Gunakan rumus chi-square untuk menghitung nilai untuk setiap term:

$$\chi^2 = \sum \left(\frac{(O_i - E_i)^2}{E_i} \right) \quad (3)$$

O_i adalah Observed frequencies

E_i adalah Expected frequencies

Contoh: Chi-Square for scarlet

$$\chi^2_{\text{scarlet}} = \left(\frac{1 - 0.3846^2}{0.3846} \right) + \left(\frac{0 - 0^2}{0} \right) + \left(\frac{0 - 0^2}{0} \right)$$

$$\chi^2_{\text{scarlet}} = 1 - 0.3846 = (0.6154)^2 = 0.3795$$

$$\chi^2_{\text{scarlet}} = \frac{0.3795}{0.3846} = 0.9867$$

Pemodelan Machine Learning

Model SVM dilatih dengan data pelatihan (80%) dan diuji dengan data pengujian (20%). Evaluasi dilakukan untuk mengukur keakuratan model dan melibatkan beberapa tahap utama dalam pemilihan parameter dan optimasi.

Pemilihan Kernel: Berbagai jenis kernel, seperti Linear, Polynomial, Radial Basis Function (RBF), dan Sigmoid, telah diuji. Penyesuaian Parameter C dan Gamma (γ): Parameter C mengatur keseimbangan antara separation margin dan classification error dengan nilai yang digunakan antara 0.1, 1, 10, 100, dan 1000. sedangkan Gamma mempengaruhi seberapa jauh pengaruh data training meluas dengan pemilihan nilai 1, 0.1, 0.01, 0.001, dan 0.0001.

Evaluasi Model

Setelah parameter dioptimalkan, dataset dibagi menjadi set pelatihan (80%) dan pengujian (20%). data uji akan digunakan untuk mengevaluasi model. Pengujian model akan dilakukan dengan menghitung metrik akurasi, presisi dan Recall serta F1-score (F-measure) berdasarkan confusion matrix. Metrik-metrik ini secara kolektif memberikan penilaian keseluruhan tentang seberapa baik kinerja model dalam membedakan antara kelas sentimen yang berbeda, contoh estimasi parameter Tabel 8.

Tabel 8. Contoh estimasi parameter

Aktual	Prediksi		
	Positif	Negatif	Netral
Positif	TP	FP	PN
Negatif	FN	TN	NNu
Netral	FNu	NuN	TNu

$$Accuracy = \frac{(TP + TN + TNu)}{TP + TN + TNu + FP + FN + FNu} \times 100 \quad (4)$$

$$Precision = \frac{TP}{TP + FN + FNu} \times 100 \quad (5)$$

$$Recall = \frac{TP}{TP + FP + PN} \times 100 \quad (6)$$

$$F - Measure = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (7)$$

- 1) True Positive (TP) adalah data positif yang diprediksi positif.
- 2) False Positive (FP) adalah data positif yang diprediksi negatif.
- 3) Positive Neutral (PN) adalah data positif yang diprediksi netral.
- 4) True Negative (TN) adalah data negatif yang diprediksi sebagai benar.
- 5) False Negative (FN) adalah data negatif yang diprediksi positif.
- 6) Negative Neutral (NNu) adalah data negatif yang diprediksi positif
- 7) True Neutral (TNu) adalah data netral yang diprediksi benar.
- 8) False Neutral (FNu) adalah data netral yang diprediksi positif
- 9) Netral Negatif (NuN) adalah data netral yang diprediksi negatif

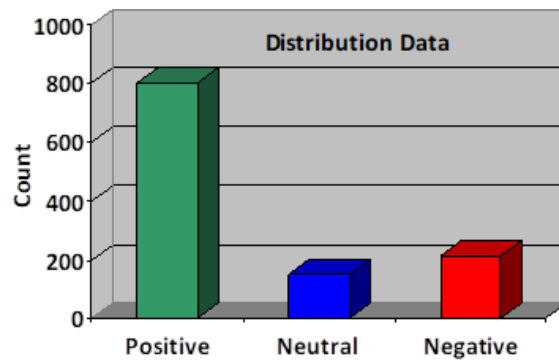
Visualisasi Data

Visualisasi data digunakan untuk memperjelas temuan dan memfasilitasi interpretasi hasil analisis sentimen. Hal ini memungkinkan pemahaman yang lebih jelas mengenai distribusi data dan menyoroti kata-kata yang sering diekspresikan dalam kumpulan data. Teknik visualisasi seperti pie chart, bar chart, dan word clouds digunakan untuk menyajikan wawasan ini

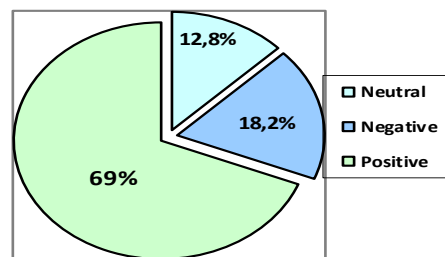
HASIL DAN PEMBAHASAN

Distribusi Sentimen

Berdasarkan analisis data sentimen, ditemukan bahwa sentimen positif mendominasi dengan persentase sebesar 69%. Hal ini menunjukkan bahwa sebagian besar opini konsumen terkait Merek X di media sosial adalah positif. (Gambar 1) dan (Gambar 2) menampilkan distribusi sentimen positif dan negatif, seperti yang diilustrasikan oleh grafik yang menyertainya.



Gambar 1. Bar chart distribusi data



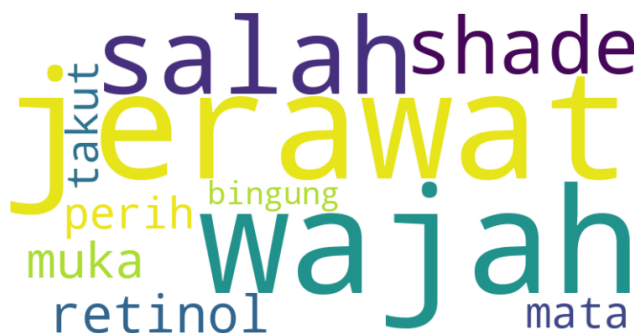
Gambar 2. Pie chart distribusi data

Faktor-Faktor yang Mempengaruhi Sentimen

Analisis mendalam mengidentifikasi beberapa faktor yang secara signifikan mempengaruhi sentimen positif dan negatif terhadap produk Merek X. Faktor-faktor ini menunjukkan pola-pola tertentu dalam tweet yang dapat memberikan wawasan penting bagi Merek X. Seperti yang dapat dilihat pada (Gambar 3) untuk sentimen positif dan (Gambar 4) untuk sentimen negatif.



Gambar 3. Wordcloud sentiment positif



Gambar 4. Wordcloud sentiment negatif

Faktor-faktor yang mempengaruhi sentimen positif:

- a. **Harga:** Banyak tweet yang berfokus pada harga suatu produk mengungkapkan seberapa puas pelanggan terhadap produk tersebut. Ulasan yang menggambarkan produk Merek X sebagai “harga yang kompetitif” dan “terjangkau” mengungkapkan bahwa pelanggan berpikir bahwa produk tersebut menawarkan nilai yang baik untuk uang yang mereka keluarkan.
- b. **Serum:** Tweet positif mengandung banyak referensi tentang “serum”. Pengguna serum yang senang dengan pembelian mereka biasanya memberikan nilai tinggi karena produk tersebut memenuhi ekspektasi mereka, seperti kulit yang lebih baik.
- c. **Kulit:** Dalam tweet, istilah “kulit” sering kali dihubungkan dengan kenangan indah. Komentar positif seperti “kulit lebih halus” atau “kulit terlihat lebih cerah” sering diberikan oleh pengguna yang melaporkan peningkatan kondisi kulit mereka.
- d. **Skincare:** Diskusi tentang rejimen “skincare” sering kali dikaitkan dengan perasaan bahagia, terutama jika produk dari Merek X dianggap penting untuk perawatan kulit mereka.
- e. **Cushion:** Kata “cushion” sering digunakan dalam konteks positif dalam evaluasi kosmetik, yang menunjukkan kebahagiaan dengan hasil produk yang menyenangkan dan memuaskan.

Faktor-faktor yang memengaruhi sentimen negatif:

- a. **Jerawat:** Tweet terkait kritik sering kali mengandung istilah yang berhubungan dengan “jerawat”. Beberapa pelanggan mengatakan bahwa produk tersebut memperburuk kondisi kulit mereka atau menimbulkan jerawat.
- b. **Wajah:** Membicarakan masalah “wajah” adalah cara tertentu untuk mengekspresikan ketidakpuasan. Pengalaman negatif biasanya dikaitkan dengan tweet yang menarik perhatian pada ketidaknyamanan atau hasil yang tidak sesuai dengan harapan.
- c. **Salah:** Ketika seseorang menggunakan istilah “salah” dalam sebuah tweet, itu berarti harapan mereka tidak terpenuhi oleh produk. Ekspresi seperti “pilihan yang salah” sering digunakan ketika sesuatu tidak berjalan dengan baik dan ketika seseorang tidak senang.
- d. **Shade:** Jika produk kosmetik yang diberikan tidak sesuai dengan yang diharapkan atau tidak sesuai dengan warna kulit, warna atau “shade” yang tidak sesuai sering kali menimbulkan perasaan yang tidak menyenangkan.
- e. **Perih:** Efek samping yang tidak diinginkan dari produk dapat menjadi penyebab ketidaknyamanan seperti “perih” yang disebutkan dalam tweet yang tidak menyenangkan.

Dengan menyadari elemen-elemen ini, Merek X dapat menentukan area yang membutuhkan pengembangan dan memodifikasi pendekatan mereka untuk meningkatkan pengalaman pelanggan. Menyelesaikan aspek-aspek yang sering dikeluhkan dapat meningkatkan kebahagiaan konsumen dan kesan merek.

Hasil Detail Parameter Model SVM

Pada penelitian ini, hasil terbaik dari optimasi parameter Support Vector Machine (SVM) adalah sebagai berikut:

- a. **Kernel:** Kernel RBF memberikan hasil terbaik dengan akurasi tertinggi,
- b. **Nilai Parameter C:** Nilai terbaik yang ditemukan adalah $C = 10$, yang memberikan keseimbangan optimal antara kompleksitas model dan akurasi.
- c. **Nilai Gamma (γ):** Nilai terbaik yang ditemukan adalah $\gamma = 1$, yang efektif dalam menangkap pola tanpa menyebabkan overfitting.

Evaluasi Model

Model diuji dengan 20% data sebagai data uji dan menggunakan confusion matrix dan didapatkan akurasi sebesar 83%. Hal ini menunjukkan bahwa parameter yang dipilih berhasil memaksimalkan kinerja model dalam mengklasifikasikan sentimen secara akurat.

Wawasan untuk Strategi Pemasaran

Hasil analisis memberikan beberapa rekomendasi strategis untuk Merek X:

- a. Memperkuat Kualitas Produk: Kualitas produk yang berkualitas tinggi penting untuk mempertahankan ulasan positif. Menanggapi umpan balik pelanggan mengenai kualitas dan melakukan perbaikan yang diperlukan dapat membantu mempertahankan reputasi positif.
- b. Memanfaatkan Umpan Balik Pelanggan: Melibatkan pelanggan dalam perbaikan produk dan secara aktif menanggapi umpan balik dapat memperkuat hubungan dan meningkatkan kepuasan.
- c. Inovasi Produk: Agar tetap relevan, Merek X harus terus berinovasi. Mengembangkan produk baru atau memperbarui produk yang sudah ada dengan fitur-fitur terbaru dapat menarik perhatian dan memenuhi harapan pelanggan.
- d. Dengan menerapkan wawasan ini, Merek X dapat meningkatkan posisi pasarnya, meningkatkan kepuasan pelanggan, dan menyempurnakan strategi pemasaran berdasarkan analisis sentimen yang mendalam.

KESIMPULAN

Penelitian ini memberikan analisis mendalam mengenai reputasi Merek X di platform Twitter melalui evaluasi sentimen. Temuan utama dari penelitian ini menunjukkan bahwa sentimen positif mendominasi dengan persentase 69%. Penggunaan model Support Vector Machine (SVM) dalam penelitian ini menghasilkan akurasi sebesar 83%, yang menandakan keefektifan model dalam klasifikasi sentimen. Faktor utama yang mempengaruhi sentimen terhadap produk Merek X diidentifikasi sebagai kualitas produk, harga, inovasi, dan kesesuaian produk berdasarkan hasil analisis. Saran strategis yang diberikan meliputi fokus pada peningkatan kualitas produk, respon proaktif terhadap umpan balik pelanggan, dan inovasi produk baru untuk memastikan daya saing di pasar. Diharapkan, insight ini dapat memandu Merek X dalam mengembangkan strategi pemasaran yang lebih tepat sasaran dan adaptif terhadap opini konsumen.

REFERENSI

- Amini, T. A., & Setiawan, K. (2024). Application of the naive bayes algorithm in twitter sentiment analysis of 2024 vice presidential candidate Gibran Rakabuming Raka using rapidminer. *International Journal Software Engineering and Computer Science*, 4(1), 234–246. <https://doi.org/10.35870/ijsecs.v4i1.2236>
- Ben-Hur, A., & Weston, J. (2010). A user's guide to support vector machines. In O. Carugo & F. Eisenhaber (Eds.), *Data Mining Techniques for the Life Sciences* (Vol. 609, pp. 223–239). Humana Press. https://doi.org/10.1007/978-1-60327-241-4_13
- Boumhidi, A., Benlahbib, A., & Nfaoui, E. H. (2022). Cross-platform reputation generation system based on aspect-based sentiment analysis. *IEEE Access*, 10, 2515–2531. <https://doi.org/10.1109/ACCESS.2021.3139956>
- Chakraborty, S., Sadhukhan, S., Bhar, A., & Ghosh, M. (2023). Sentiment analysis in the perspective of natural language processing. *International Journal for Research in Applied Science and Engineering Technology*, 11(11), 2235–2241. <https://doi.org/10.22214/ijraset.2023.56925>
- Cortes, C., & Vapnik, V. (1995). Support-vector networks editor. *Machine Learning*, 20(3), 273–297. <https://doi.org/10.1007/BF00994018>

- Dami, S., & Alimardani, R. (2024). An aspect-level sentiment analysis based on lda topic modeling. *Journal of Information Systems and Telecommunication*, 12(2), 117–126. <https://doi.org/10.61186/jist.38104.12.46.117>
- Das, M., Kamalanathan, S., & Alphonse, P. J. A. (2021). A comparative study on TF-IDF feature weighting method and its analysis using unstructured dataset. *COLINS 2021: Proceedings of the 5th International Conference. April 22-23, 2021.*, 2870, 98–107. <https://doi.org/10.48550/arXiv.2308.04037>
- Dubey, T., & Jain, A. (2019). Sentiment analysis of keenly intellective smart phone product review utilizing svm classification technique. *ICCCNT 2019: Proceedings of the 10th International Conference on Computing, Communication and Networking Technologies. July 06-08, 2019.*, 1–8. <https://doi.org/10.1109/ICCCNT45670.2019.8944795>
- Fadlil, A., Riadi, I., & Andrianto, F. (2024). Improving sentiment analysis in digital marketplaces through svm kernel fine-tuning. *International Journal of Computing and Digital Systems*, 16(1), 159–171. <https://doi.org/10.12785/ijcds/160113>
- Fatmawati, I., & Fauzan, N. (2021). Building customer trust through corporate social responsibility: the effects of corporate reputation and word of mouth. *Journal of Asian Finance, Economics and Business*, 8(3), 0793–0805. <https://doi.org/10.13106/jafeb.2021.vol8.no3.0793>
- Fournier, S., & Srinivasan, S. (2023). Monitoring marketing sources of brand reputation risk. *NIM Marketing Intelligence Review*, 15(1), 46–53. <https://doi.org/10.2478/nimmir-2023-0007>
- Ganga, B. M., Meghana, K. S., Raykar, S. S., Spandana, M. K., & Vinutha, V. (2022). Sentimental analysis on cosmetics using machine learning. *ACS Journal for Science and Engineering*, 2(2), 11–29. <https://doi.org/10.34293/acsjse.v2i2.31>
- Imanuddin, S. H., Adi, K., & Gernowo, R. (2023). Sentiment analysis on satuschat application using support vector machine method. *Journal of Electronics, Electromedical Engineering, and Medical Informatics*, 5(3), 143–149. <https://doi.org/10.35882/jeemi.v5i3.304>
- Jain, R., Rai, R. S., Jain, S., Ahluwalia, R., & Gupta, J. (2023). Real time sentiment analysis of natural language using multimedia input. *Multimedia Tools and Applications*, 82(26), 41021–41036. <https://doi.org/10.1007/s11042-023-15213-3>
- Jannah, N. Z. B., & Kusnawi, K. (2024). Comparison of naïve bayes and svm in sentiment analysis of product reviews on marketplaces. *Sinkron*, 8(2), 727–733. <https://doi.org/10.33395/sinkron.v8i2.13559>
- Kavyasri, G., Keerthana, D., Keerthana, K., Reddy, K. B., Keerthi, K., Aditya, K. J., & Arivazhagan, D. (2023). Margin maximization of text classification based on support vector machine. *International Journal for Research in Applied Science & Engineering Technology*, 11(12), 2321–9653. <https://doi.org/10.22214/ijraset.2023.57420>
- Kayakuş, M., Açıkgoz, F. Y., Dinca, M. N., & Kabas, O. (2024). Sustainable brand reputation: evaluation of iPhone customer reviews with machine learning and sentiment analysis. *Sustainability*, 16(14), 6121. <https://doi.org/10.3390/su16146121>
- Khuspe, A., Gaikwad, T., Sarkar, A., Wyawahare, M., Kumari, A., & Chopde, A. (2024). Sentiment analysis on twitter through machine learning: a comprehensive approach with user-centric visualisations. *International Journal of Innovative Science and Research Technology*, 9(4), 2066–2073. <https://doi.org/10.38124/ijisrt/ijisrt24apr227>
- Lin, L. L.-C., Yang, G.-H., Wang, Y.-T., Anggarina, P. T., & Irawan, A. P. (2023). The impact of brand reputation and marketing tools on purchasing intentions -care products as an example. *International Journal of Application on Economics and Business*, 1(3), 1557–1564. <https://doi.org/10.24912/ijaeb.v1i3.1557-1564>

- Maiti, A., Abarda, A., & Hanini, M. (2024). The impact of feature extraction techniques on the performance of text data classification models. *Indonesian Journal of Electrical Engineering and Computer Science*, 35(2), 1041–1052. <https://doi.org/10.11591/ijeecs.v35.i2.pp1041-1052>
- Makhtum, A. R., & Muhajir, M. (2023). Sentiment analysis of omnibus law using support vector machine (svm) with linear kernel. *BAREKENG: Jurnal Ilmu Matematika Dan Terapan*, 17(4), 2197–2206. <https://doi.org/10.30598/barekengvol17iss4pp2197-2206>
- Mamuaya, N. C. (2024). Investigating the impact of product quality, price sensitivity, and brand reputation on consumer purchase intentions in the FMCG sector. *International Journal of Business, Law, and Education*, 5(2), 1576–1583. <https://doi.org/10.56442/ijble.v5i2.614>
- Manimegalai, D., & Srinivas, V. V. (2023). Sentiment analysis on the brand recall of twitter data with the tool of NCSU tweet sentiment visualization. *International Journal of Research in Marketing Management and Sales*, 5(1), 48–53. <https://doi.org/10.33545/26633329.2023.v5.i1a.121>
- Mariutti, F. G., & Giraldi, J. de M. E. (2020). Country brand equity: the role of image and reputation. *BAR - Brazilian Administration Review*, 17(3), 1–24. <https://doi.org/10.1590/1807-7692bar2020180128>
- Mohammad, R., Movafegh, A., & Fotoohi, H. (2015). Brand reputation and costumer trust. *Arabian Journal of Business and Management Review*, 5(2), 48–52. <https://doi.org/10.12816/0019012>
- Mostaghel, R., Oghazi, P., Patel, P. C., Parida, V., & Hultman, M. (2019). Marketing and supply chain coordination and intelligence quality: A product innovation performance perspective. *Journal of Business Research*, 101, 597–606. <https://doi.org/10.1016/j.jbusres.2019.02.058>
- Mutia, A. (2022, December 5). *Beauty and personal care product revenue in indonesia to reach IDR111.83 trillion by 2022*. databoks.katadata.co.id. <https://databoks.katadata.co.id/datapublish/2022/12/05/makin-meroket-pendapatan-produk-kecantikan-dan-perawatan-diri-di-ri-capai-rp11183-triliun-pada-2022>
- Olaleye, S. A., Sanusi, I. T., & Salo, J. (2018). Sentiment analysis of social commerce: A harbinger of online reputation management. *International Journal of Electronic Business*, 14(2), 85–102. <https://doi.org/10.1504/IJEB.2018.094864>
- Purba, J. T., Budiono, S., & Rajagukguk, W. (2022, April 5). Performance of collaboration social media marketing and product quality on local brand loyalty of coffee shop for revisit intention. *Paper Presented at the 3rd African International Conference. Industrial Engineering and Operations Management. April 5-7, 2022*. <https://doi.org/10.46254/AF03.20220081>
- Qaiser, S., Yusoff, N., Ali, R., Remli, M. A., & Adli, H. K. (2021). A comparison of machine learning techniques for sentiment analysis. *Turkish Journal of Computer and Mathematics Education*, 12(3), 1738–1744. <https://doi.org/10.17762/turcomat.v12i3.999>
- Rahmawati, Fuadi, W., & Afrillia, Y. (2024). Cosmetic shop sentiment analysis on tiktok shop using the support vector machine method. *International Journal of Engineering, Science & Information Technology*, 4(2), 31–37. <https://doi.org/10.52088/ijesty.v4i1.498>
- Rimpy, Dhankhar, A., & Dhankhar, A. (2024). Sentimental analysis of social networks: A comprehensive review (2018-2023). *Multidisciplinary Reviews*, 7(7), 2024126. <https://doi.org/10.31893/multirev.2024126>
- Saragih, H., & Manurung, J. (2024). Leveraging the BERT model for enhanced sentiment analysis in multicontextual social media content. *Jurnal Teknik Informatika C.I.T Medicom*, 16(2), 82–89. <https://doi.org/10.35335/cit.vol16.2024.766.pp82-89>

- Singh, A., & Banerjee, S. (2024). A review of: ensemble feature selection scheme-based performance evaluation of several classifiers for sentiment analysis. *International Journal for Research in Applied Science and Engineering Technology*, 12(6), 1833–1840. <https://doi.org/10.22214/ijraset.2024.63404>
- Singh, N., & Jaiswal, U. C. (2023). A detailed sentiment analysis survey based on machine learning techniques. *Advances in Distributed Computing and Artificial Intelligence Journal*, 12(1), e29105. <https://doi.org/10.14201/adcaij.29105>
- Starbuck, C. (2023). Data preparation. In *The Fundamentals of People Analytics* (1st ed., pp. 79–95). Springer International Publishing. https://doi.org/10.1007/978-3-031-28674-2_6
- statista. (2023, October). *Skin care: market data & analysis. Skin care - Indonesia*. Statista.Com. <https://www.statista.com/outlook/cmo/beauty-personal-care/skin-care/indonesia>
- Stravinskiene, J., Matuleviciene, M., & Hopeniene, R. (2021). Impact of corporate reputation dimensions on consumer trust. *Inzinerine Ekonomika-Engineering Economics*, 32(2), 177–192. <https://doi.org/10.5755/j01.ee.32.2.27548>
- Sudhakaran, P., & Jaiganesh, M. (2020). Sentiment analysis based product selection for enhancing e-commerce. *International Journal of Innovative Technology and Exploring Engineering*, 9(3S), 389–394. <https://doi.org/10.35940/ijitee.c1083.0193s20>
- Suryawanshi, N. S. (2024). Sentiment analysis with machine learning and deep learning: A survey of techniques and applications. *International Journal of Science and Research Archive*, 12(2), 005–015. <https://doi.org/10.30574/ijrsra.2024.12.2.1205>
- Tellawar, A., Wattamwar, S., & Thorat, B. (2023). Online fake review detection using svm. *International Journal for Research in Applied Science and Engineering Technology*, 11(5), 2006–2008. <https://doi.org/10.22214/ijraset.2023.51925>
- Tembhurne, J. V., Agrawal, A., & Lakhota, K. (2024). Covid-19 twitter sentiment classification using hybrid deep learning model based on grid search methodology. *ArXiv*, 2406.10266. <https://doi.org/10.48550/arxiv.2406.10266>
- Traore, N. (2024). Impact of online reviews on brand reputation and customer purchase behaviors in online retail market. *International Journal of Innovative Research in Multidisciplinary Education*, 03(03), 381–383. <https://doi.org/10.58806/ijirme.2024.v3i3n17>
- Vaibhavi, S. S., Devadathan, R., Inbamala, T. A., Susmita, P., Supritha, R., Sanjayan, J., & Bhuvana, J. (2024). A Study And Development Of Application On Sentiment Analysis. *International Scientific Journal of Engineering and Management*, 03(03), 1–7. <https://doi.org/10.55041/isjem01354>
- Waluyo, D., Nuraini, R., & Sari, E. I. (2023, November 3). *Beautiful phenomenon of the cosmetic industry*. Indonesia.Go.Id. <https://indonesia.go.id/kategori/editorial/7804/beautiful-phenomenon-of-the-cosmetic-industry?lang=2>
- Xi, R. (2024). A comprehensive review of text sentiment analysis: a survey of traditional methods and deep learning approaches. *Science and Technology of Engineering, Chemistry and Environmental Protection*, 1(5), TE000412. <https://doi.org/10.61173/7wthqa25>
- Xiao, H., & Luo, L. (2024). An automatic sentiment analysis method for short texts based on transformer-BERT hybrid model. *IEEE Access*, 12, 93305–93317. <https://doi.org/10.1109/ACCESS.2024.3422268>
- ZAP Clinic, & MarkPlus. (2024, January 30). *Zap beauty index 2024 dan men/o/logy index*. Zapclinic.Com. https://zapclinic.com/files/ZAP_Beauty_Index_2024.pdf

- Zhan, T., Shi, C., Shi, Y., Li, H., & Lin, Y. (2024). Optimization techniques for sentiment analysis based on LLM (GPT-3). *Applied and Computational Engineering*, 77, 252–257. <https://doi.org/10.54254/2755-2721/67/2024MA0060>
- Zheng, D. (2024). Sentiment analysis for film reviews based on random forest. *Science and Technology of Engineering, Chemistry and Environmental Protection*, 1(7), TE001562. <https://doi.org/10.61173/5t8epb44>