

Analisis Sentimen Terhadap Produk *Sunscreen* Pada *Marketplace* Menggunakan *Support Vector Machine* (SVM)

Thoriq Bahtiar, Badie'ah, Imam Much Ibnu Subroto

Teknik Informatika, Fakultas Teknologi Industri, Universitas Islam Sultan Agung

Correspondence Author: thoriqbahtiar23@std.unissula.ac.id

Abstrak

Pasar online seperti Shopee telah menjadi platform penting bagi bisnis untuk menjual produk dan layanan kepada khalayak luas, mendorong pertumbuhan platform marketplace. Ulasan konsumen tidak hanya mencerminkan pengalaman pribadi pengguna dengan sebuah produk, tetapi juga memberikan informasi detail tentang efektivitas produk tersebut. Dengan menganalisis ulasan ini, calon pembeli bisa mendapatkan gambaran yang lebih jelas mengenai kelebihan dan kekurangan produk tersebut, sehingga bisa menjadi umpan balik yang berharga untuk peningkatan produk di masa mendatang. Penelitian ini bertujuan untuk mengembangkan sistem analisis sentimen terhadap produk *sunscreen* di marketplace Shopee menggunakan metode *Support Vector Machine* (SVM). Analisis sentimen dilakukan untuk mengevaluasi tingkat kepuasan konsumen terhadap produk *sunscreen* yang dijual di Shopee. Melalui analisis ini, persepsi konsumen tentang produk tersebut dapat dievaluasi, sehingga memberikan wawasan berharga untuk peningkatan produk dan pengambilan keputusan konsumen. Proses analisis sentimen mencakup tahapan pra-pemrosesan teks, termasuk *lowercasing*, *tokenization*, dan penghapusan *stop words* menggunakan *Sastrawi*. Berdasarkan hasil penelitian, metode SVM dengan pembobotan *TF-IDF* menunjukkan performa yang memuaskan dengan akurasi tinggi, nilai *precision* 0,89, *recall* 0,90, dan *F1-score* 0,89, yang mengindikasikan kemampuan model dalam mengklasifikasikan sentimen konsumen secara efektif dan konsisten.

Keyword: Analisis Sentimen, *Support Vector Machine*, Pra-pemrosesan teks

1. PENDAHULUAN

Pasar online seperti Shopee telah menjadi platform penting bagi bisnis untuk menjual produk kepada khalayak luas. Dengan munculnya pasar digital, cara konsumen berbelanja berubah secara drastis. Marketplace seperti Shopee menyediakan ruang virtual di mana penjual dapat menjangkau pelanggan potensial secara global [1]. Salah satu aspek penting dari pasar digital ini adalah memahami persepsi konsumen terhadap produk yang dijual, yang dapat dilakukan melalui analisis sentimen. Dalam penelitian ini, dilakukan analisis sentimen terhadap ulasan konsumen untuk produk *sunscreen* di Shopee, dengan menggunakan metode *Support Vector Machine* (SVM) [2].

Tujuan tugas akhir ini adalah untuk menganalisis sentimen konsumen dengan cara mengevaluasi tanggapan mereka terhadap beberapa produk *sunscreen* badan pada marketplace Shopee, baik dari segi kualitas, harga, maupun pengalaman penggunaan, sehingga dapat memberikan wawasan yang lebih mendalam mengenai preferensi, kepuasan, dan kekecewaan pelanggan. Dalam upaya meningkatkan efektivitas perbaikan atau pengembangan produk, penulis melakukan analisis mendalam terhadap sentimen pelanggan, yang memungkinkan untuk mengidentifikasi area spesifik yang membutuhkan perhatian, sehingga perbaikan atau inovasi yang diimplementasikan dapat lebih tepat sasaran dan sesuai dengan kebutuhan serta harapan pasar. Dengan menyediakan informasi yang komprehensif dan berharga, dapat membantu konsumen dalam proses pengambilan keputusan yang lebih baik dan lebih informatif, sehingga mereka dapat memilih produk atau layanan yang paling sesuai dengan kebutuhan dan preferensi mereka, serta merasa lebih percaya diri dengan pilihan yang dibuat. Dengan menggunakan sistem analisis data, dapat melihat dan memahami secara mendalam persepsi konsumen terhadap berbagai produk *sunscreen* yang dijual di platform *e-commerce* Shopee yang diberikan oleh pengguna, sehingga dapat memperoleh wawasan yang berharga mengenai kepuasan, kekhawatiran dan preferensi mereka terhadap produk-produk tersebut.

Analisis sentimen merupakan metode analisis data yang digunakan untuk menentukan opini, sentimen, atau emosi yang terkandung dalam teks [3]. Penerapan analisis sentimen telah meluas ke berbagai bidang, termasuk pemasaran, layanan pelanggan, dan riset pasar. Dengan berkembangnya *e-commerce*, analisis sentimen memainkan peran penting dalam memahami persepsi konsumen terhadap produk yang dijual di

platform seperti Shopee. Salah satu metode yang efektif untuk analisis sentimen adalah *Support Vector Machine* (SVM) [4]. Sistem analisis sentimen terhadap produk *sunscreen* di Shopee [5] menggunakan *Support Vector Machine* (SVM) adalah pendekatan yang efektif untuk memahami opini konsumen. Dengan analisis sentimen, penjual dan produsen dapat memperoleh wawasan yang mendalam tentang ulasan konsumen dan melakukan perbaikan atau pengembangan produk yang lebih tepat sasaran [6]. Implementasi SVM dalam analisis sentimen [7] menawarkan akurasi yang tinggi dan mampu menangani data berdimensi tinggi, menjadikannya pilihan yang tepat untuk melakukan penelitian ini.

Pengujian analisis sentimen pada produk kecantikan. Proses *preprocessing* yang meliputi penghapusan tanda baca, normalisasi, pembersihan data, dan *stemming* tanpa *stopword*. Dengan ekstraksi fitur TF-IDF, menggunakan Kernel Sigmoid pada SVM menunjukkan performa yang lebih baik dibandingkan kernel linear dan RBF, dengan mencapai nilai akurasi tertinggi sebesar 85,98%. Hal ini menunjukkan bahwa pemilihan kernel dapat berpengaruh signifikan terhadap hasil akhir dalam analisis sentimen ini [8].

Dengan menggunakan metode analisis sentimen menggunakan *Support Vector Machine* (SVM) dan *Naïve Bayes Classifier* (NBC), evaluasi terhadap ulasan produk The Tea Tree Skin Clearing Toner dari The Body Shop dilakukan. Hasil menunjukkan bahwa SVM dengan kernel linear memberikan tingkat akurasi sebesar 86% dengan nilai Area Under Curve (AUC) sebesar 0,91, yang lebih tinggi dibandingkan dengan metode Naïve Bayes yang memberikan akurasi sebesar 83% dengan AUC 0,82. Oleh karena itu, SVM dengan kernel linear direkomendasikan untuk klasifikasi ulasan produk ini karena kinerja yang lebih baik dalam mengidentifikasi sentimen positif, negatif, dan netral [9].

Diperoleh bahwa hasil performansi terbaik untuk klasifikasi *sentiment analysis* pada *review* buku novel berbahasa Inggris, yaitu pada penggunaan Kernel Gaussian RBF untuk setiap kedua pembobotan fitur dengan seluruh nilai persentase seleksi fitur yang digunakan dengan nilai performansi terbaik sebesar 74.2%. Selain itu, dari hasil penelitian pada kedua proses pembobotan fitur dalam satu seleksi fitur serta penggunaan kernel Linear pada klasifikasi *Support Vector Machine* (SVM), diperoleh bahwa hasil performansi yang diperoleh sebesar 70,7% jika menggunakan pembobotan fitur *Term Frequency-Inverse Document Frequency* (TF-IDF) dan seleksi fitur Chi Square. Sedangkan, hasil performansi yang berhasil diperoleh sebesar 71% jika menggunakan pembobotan fitur *Term Frequency* (TF) dan seleksi fitur Chi Square [10].

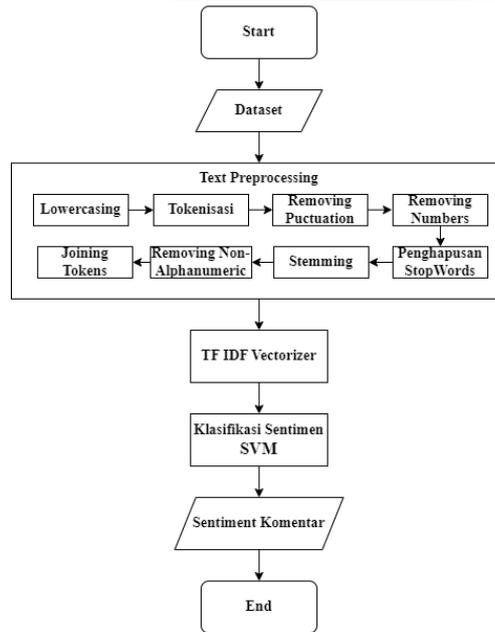
Performa klasifikasi yang dihasilkan menunjukkan nilai G-mean dan AUC terbaik untuk data uji Bukalapak dengan masing-masing sebesar 0,85 dan 0,86 pada fold pertama. Sementara itu, untuk data uji Shopee, nilai G-mean dan AUC terbaik tercatat sebesar 0,76 dan 0,77 pada fold ketujuh. Adapun data uji Tokopedia menunjukkan performa yang optimal dengan nilai G-mean sebesar 0,82 dan AUC sebesar 0,83 pada fold keenam [11].

Berdasarkan hasil pengujian yang telah dilakukan [2], ditemukan bahwa skenario 3 dan skenario 8 memiliki performansi sistem yang lebih baik dalam mengklasifikasikan tweet dibandingkan dengan skenario lainnya. Pada skenario 3, yang menggunakan kombinasi TF-IDF dan *stemming*, sistem mencapai akurasi sebesar 81,58%. Sementara itu, pada skenario 8, yang menggunakan kombinasi *word count* dan *stemming*, akurasi yang dicapai adalah 77,56%.

Dari beberapa hasil penelitian diatas peneliti dapat mengambil kesimpulan untuk membuat sebuah sistem yang mengambil data dari sebuah *platform marketplace* yang nantinya diolah untuk dilakukan analisis sentimen berdasarkan komentar yang diberikan oleh pengguna.

2. METODE PENELITIAN

Dalam penelitian ini, metode atau algoritma yang digunakan adalah *Support Vector Machine* (SVM) dan *Term Frequency-Inverse Document Frequency* (TF-IDF). Kombinasi kedua metode ini mencakup SVM sebagai teknik klasifikasi dan TF-IDF sebagai metode penghitungan bobot kata untuk *dataset*. Dengan kombinasi ini, dihasilkan sebuah sistem analisis sentimen yang efisien. *Support Vector Machine* (SVM) adalah teknik pembelajaran dengan banyak kualitas yang diinginkan dan menjadikan algoritma SVM sangat populer. SVM mempunyai dasar teoritis yang kuat dan melakukan klasifikasi lebih akurat daripada kebanyakan algoritma lain di banyak aplikasi. Banyak penelitian telah melaporkan bahwa SVM merupakan metode yang paling akurat untuk klasifikasi teks. SVM juga banyak digunakan dalam klasifikasi sentimen Adapun tahapan yang harus dilakukan dalam penelitian ini, antara lain : *Web Scraping* [12], *Text Preprocessing* [13], TF-IDF *vectorizer* [14], klasifikasi sentimen SVM dan hasil analisa sistem yaitu *confusion matrix* [15]. Gambar 1 merupakan *flowchart* rancangan model yang dibangun :



Gambar 1. Flowchart Alur Sistem

- a. *Start* : Langkah awal dari proses.
- b. *Dataset* : *Dataset* merupakan hasil *scraping* komentar produk Shopee yang akan digunakan dalam analisis sentimen.
- c. *Text Preprocessing* : Langkah-langkah pemrosesan teks menjadi lebih bersih, untuk menghasilkan sebuah sistem dengan akurasi yang tinggi.
- d. TF-IDF Vectorizer : Mengubah teks yang sudah diproses menjadi vektor angka menggunakan metode TF-IDF (Term Frequency-Inverse Document Frequency) untuk mempresentasikan bobot pentingnya kata dalam suatu dokumen.
- e. Klasifikasi Sentimen (SVM) : Menggunakan algoritma *Support Vector Machine* (SVM) untuk mengklasifikasikan sentimen teks berdasarkan vektor yang sudah dihasilkan dari TF-IDF.
- f. Sentimen Komentar : Hasil dari klasifikasi sentimen yang menunjukkan apakah sentimen komentar adalah positif atau negatif.
- g. *End* : Langkah akhir dari proses.

Metode data *preprocessing* merupakan langkah-langkah yang penting dalam analisis sentimen untuk mempersiapkan data teks agar sesuai untuk analisis. Beberapa contoh untuk proses data *preprocessing* :

Tabel 1. Hasil *Lowercasing*

Sebelum	Sesudah
Packingan nya tebal bgt padahal cmn beli 1, pengemasan cepat, sesuai pesanan, exp masi 2025	packingan nya tebal bgt padahal cmn beli 1, pengemasan cepat, sesuai pesanan, exp masi 2025

Tabel 2. Hasil Tokenisasi

Sebelum	Sesudah
Nyoba produk baru, kemasan baru, praktis dan mungil. Pengiriman cepat. Semoga cocok di kulit aku.	'Nyoba', 'produk', 'baru', ',', 'kemasan', 'baru', ',', 'praktis', 'dan', 'mungil', ',', 'Pengiriman', 'cepat', ',', 'Semoga', 'cocok', 'di', 'kulit', 'aku', ','

Tabel 3. Hasil *Removing Punctuation*

Sebelum	Sesudah
oke banget packaging super aman suka! and over all very well done thank u yaa!	oke banget packaging super aman suka and over all very well done thank u yaa

Tabel 4. Hasil *Removing Numbers*

Sebelum	Sesudah
Packingan nya tebal bgt padahal cmn beli 1, pengemasan cepat, sesuai pesanan, exp masi 2025	Packingan nya tebal bgt padahal cmn beli, pengemasan cepat, sesuai pesanan, exp masi

Tabel 5. Hasil Penghapusan *Stopword*

Sebelum	Sesudah
Smp udh bner ² abs pke nya.. Enk bgt d pke, hasilnya mate tpi gaa berat. Kyk pke bedak aja. Wangi nya jg enk kyk bedak bayi	Smp udh bner ² abs pke Enk bgt pke hasilnya mate berat Kyk pke bedak Wangi enk kyk bedak bayi

Tabel 6. Hasil *Stemming*

Sebelum	Sesudah
Bagus banget cocok untuk kulit remaja untuk kulit yang kusam harganya terjangkau	Bagus banget cocok untuk kulit remaja untuk kulit yang kusam harga jangkau

Tabel 7. Hasil *Removing Non-Alphanumeric*

Sebelum	Sesudah
Pengirimannya cepet banget, ga sampai 3 hari udah nyampe 🤗👍 ini 100% ori ya gaisss jadi jangan khawatir, packingnya pake kardus dan aman bgt, kurirnya ramah karena udah langganan, thanks eminaa 🥰❤️	Pengirimannya cepet banget ga sampai 3 hari udah nyampe ini 100 ori ya gaisss jadi jangan khawatir packingnya pake kardus dan aman bgt kurirnya ramah karena udah langganan thanks eminaa

Tabel 8. Hasil *Joining Tokens*

Sebelum	Sesudah
'Nyoba', 'produk', 'baru', ',', 'kemasan', 'baru', ',', 'praktis', 'dan', 'mungil', ',', 'Pengiriman', 'cepat', ',', 'Semoga', 'cocok', 'di', 'kulit', 'aku', ','	Nyoba produk baru , kemasan baru , praktis dan mungil . Pengiriman cepat . Semoga cocok di kulit aku .

Pada proses *confusion matrix* ini adalah langkah krusial dalam pembangunan model pembelajaran mesin, termasuk model *Support Vector Machine* (SVM). Evaluasi hasil bertujuan untuk menilai kinerja model yang telah dilatih dengan menggunakan data pengujian (*test dataset*) yang tidak digunakan selama pelatihan [15]. Akurasi :

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

Precision :

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

Recall :

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

F1 Score :

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

Keterangan :

TP = *True Positif*. Adalah jumlah contoh di mana model benar-benar memprediksi kelas positif dan kelas sebenarnya juga positif. Ini berarti model berhasil mengidentifikasi contoh yang benar-benar positif.

TN = *True Negatif*. Adalah jumlah contoh di mana model benar-benar memprediksi kelas negatif dan kelas sebenarnya juga negatif. Ini berarti model berhasil mengidentifikasi contoh yang benar-benar negatif.

FP = *False Positive*. Adalah jumlah contoh di mana model memprediksi kelas positif padahal kelas sebenarnya adalah negatif. Ini berarti model salah dalam memprediksi contoh sebagai positif ketika seharusnya negatif.

FN = *False Negative*. Adalah jumlah contoh di mana model memprediksi kelas negatif padahal kelas sebenarnya adalah positif. Ini berarti model salah dalam memprediksi contoh sebagai negatif ketika seharusnya positif.

3. HASIL DAN ANALISA

Pada proses penelitian ini yaitu dengan mengumpulkan ulasan dari Shopee, kemudian melakukan pra-pemrosesan untuk membersihkan dan menyiapkan data, melatih model SVM untuk mengklasifikasikan sentimen dan kemudian mengevaluasi serta mengimplementasikan model untuk menganalisis ulasan baru.

3.1 Web Scraping

Proses pengambilan *dataset* dilakukan menggunakan *software* WebHarvy, sebuah *software* yang efisien untuk *web scraping*. Dalam penelitian ini, data yang diambil berjumlah total 636 *dataset*, yang masing-masing terdiri dari ulasan-ulasan produk *sunscreen* dari Shopee. *Dataset* ini diorganisasikan dengan rinci, di mana setiap merk mendapatkan 106 *dataset*. Merk-merk yang dianalisis meliputi Make Over, Wardah, Emina, Skintific, Garnier, dan CosRX. Tabel 3.1 merupakan beberapa contoh hasil yang diperoleh dari proses *web scraping* :

Tabel 9. Hasil *Web Scraping*

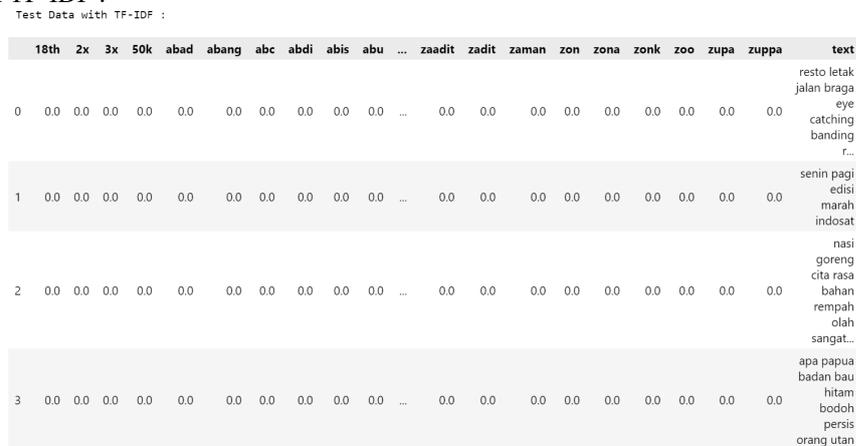
Merk	Ulasan
MakeOver	Lamaaaa bgt packing dan pengirimannya. Aman dan lengkap seluruh produk pesanannya, tanggal expired nya aman. Untuk bonusnya itu dipakenya gimana yah sama powdernya aja beda ukuran tempatnya hehe
Wardah	Pengiriman cepat, harga yang terjangkau dan produk sesuai dengan deskripsi 🍑 Aku sih ngerasanya ini ringan bngt dipakai cepet nyerep gtu jdi gk bikin wajah jdi white cast luv deh 💖
Emina	Pengiriman nya rapih dan respon sangat cepat. Rekomendasi deh beli di sini. Lain kali beli lagi. Semoga tetap menjadi toko yang amanah
Skintific	Dari kemaren pengen nyoba, akhirnya nyoba juga. Wangii yaa ternyata, cocok dibawa kemana aja nih simple tinggal semprot. Packingnya amann bgt tebal bublewarp nyaa 😊😊
Garnier	Profil Kecantikan: jenis kulit kering dan kusam Kemasan: keren Manfaat: belum di coba Semoga cocok yaa..
CosRX	Produknya blm saya coba ke wajah langsung, cuman pas saya coba ke tangan wangi terus lembut juga. Thank you!!! Pengirimannya sama pengemasannya cepet banget 🥰🥰👍👍

3.2 Text Preprocessing

Setelah berhasil mendapatkan *dataset* ulasan komentar dari *platform e-commerce* Shopee, langkah selanjutnya adalah melakukan serangkaian proses pemrosesan teks untuk membersihkan data tersebut. Proses ini meliputi *lowercasing*, tokenisasi, *removing punctuation*, *removing numbers*, penghapusan *stop words*, *stemming*, *removing non-alphanumeric*, *joining tokens*. Semua langkah ini bertujuan untuk mengubah teks mentah menjadi data yang lebih terstruktur dan siap untuk dianalisis, sehingga proses analisis sentimen dapat dilakukan dengan lebih akurat dan efisien.

3.3 TF-IDF

Pada proses perhitungan TF-IDF, yang merupakan teknik penting dalam pemrosesan teks dan analisis informasi. Teknik ini bertujuan untuk mengevaluasi seberapa penting sebuah kata dalam dokumen relatif terhadap kumpulan dokumen yang lebih besar, dengan mempertimbangkan frekuensi kemunculan kata dalam dokumen serta seberapa jarang kata tersebut muncul di seluruh dokumen. Kode program ini menjalankan berbagai langkah, mulai dari tokenisasi teks hingga perhitungan dan normalisasi nilai TF-IDF, untuk menghasilkan representasi yang berguna bagi analisis teks dan pencarian informasi. Gambar 2 merupakan hasil perhitungan TF-IDF :



Gambar 2. Hasil TF-IDF

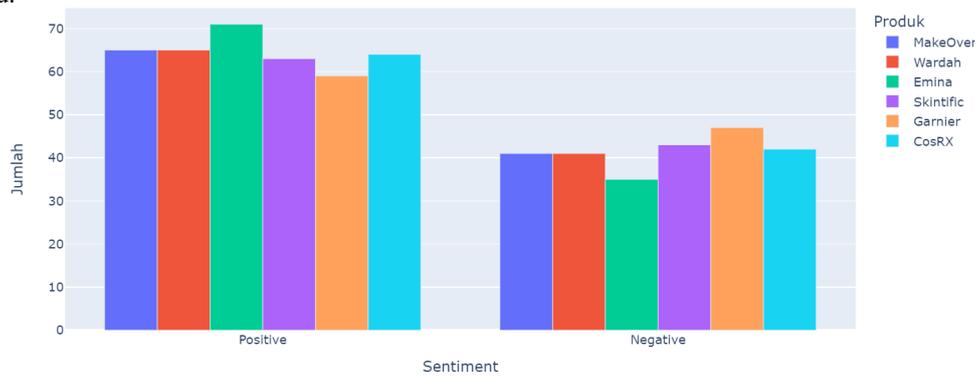
3.4 Support Vector Machine (SVM)

Pada proses pelatihan model *Support Vector Machine (SVM)*, yang merupakan metode pembelajaran mesin yang populer untuk klasifikasi dan regresi, metode ini digunakan untuk mengimplementasikan berbagai

langkah kritis dalam pelatihan model. Proses ini melibatkan pengolahan data *input*, penentuan parameter model, pelatihan model menggunakan *dataset* yang telah disiapkan, serta evaluasi kinerja model berdasarkan metrik yang relevan. Kode ini mencakup tahapan seperti pemisahan data menjadi *dataset* pelatihan dan *dataset* pengujian, penerapan teknik normalisasi atau standarisasi data jika diperlukan, serta optimasi *hyperparameter* untuk mencapai performa terbaik dari model SVM yang dihasilkan.

3.5 Hasil Penelitian

Berdasarkan hasil analisis sentimen yang sudah dilakukan terhadap berbagai ulasan pengguna, di mana kami mengidentifikasi pola-pola emosional yang dominan, baik yang bersifat positif maupun negatif, untuk memberikan wawasan yang lebih mendalam mengenai persepsi umum terhadap produk atau layanan yang ditinjau.



Gambar 3. Hasil Penelitian

- MakeOver :
Sentimen Positif : 65
Sentimen Negatif : 41
- Wardah :
Sentimen Positif : 65
Sentimen Negatif : 41
- Emina :
Sentimen Positif : 71
Sentimen Negatif : 35
- Skintific :
Sentimen Positif : 63
Sentimen Negatif : 43
- Garnier :
Sentimen Positif : 58
Sentimen Negatif : 48
- CosRX
Sentimen Positif : 64
Sentimen Negatif : 42

Produk dengan Sentimen Positif Tertinggi : Emina memiliki jumlah sentimen positif tertinggi, mencapai sebanyak 71. Produk dengan Sentimen Negatif Tertinggi : Garnier memiliki jumlah sentimen negatif tertinggi, mencapai sebanyak 48.

Secara keseluruhan, grafik ini menunjukkan bahwa mayoritas produk memiliki ulasan positif yang lebih banyak daripada ulasan negatif, dengan Emina sebagai produk yang paling disukai berdasarkan jumlah sentimen positif yang diterima. Garnier, meskipun memiliki jumlah sentimen positif yang cukup tinggi, juga memiliki sentimen negatif yang cukup signifikan, menunjukkan adanya beberapa ketidakpuasan di antara pengguna.

3.6 Hasil Confusion Matrix

Hasil evaluasi yang mencakup analisis kinerja model pada data *test* serta perbandingannya dengan kinerja pada data *train* (data pelatihan), memberikan wawasan mendalam mengenai efektivitas model dalam memprediksi label target.

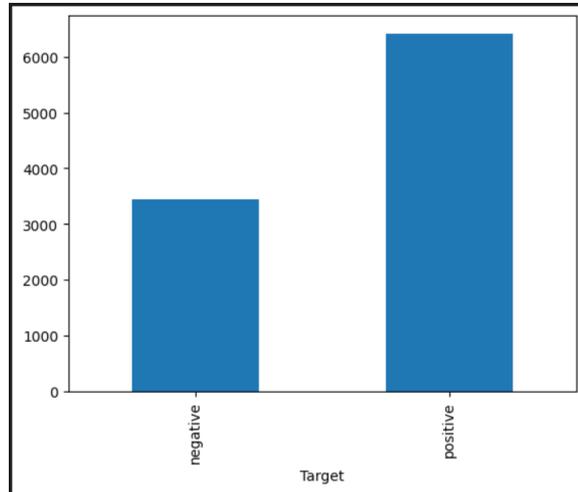
Tabel 10. *Confusion Matrix*

	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>	<i>Support</i>
<i>Positive</i>	0,89	0,90	0,89	735
<i>Negative</i>	0,90	0,89	0,89	735
<i>Accuracy</i>			0,89	1470
<i>Macro Avg</i>	0,89	0,89	0,89	1470
<i>Weighted Avg</i>	0,89	0,89	0,89	1470

Secara keseluruhan, hasil evaluasi model memiliki performa yang baik dengan *accuracy* sebesar 0.89, serta *precision*, *recall* dan *f1-score* yang relatif seimbang antara kelas positif dan negatif.

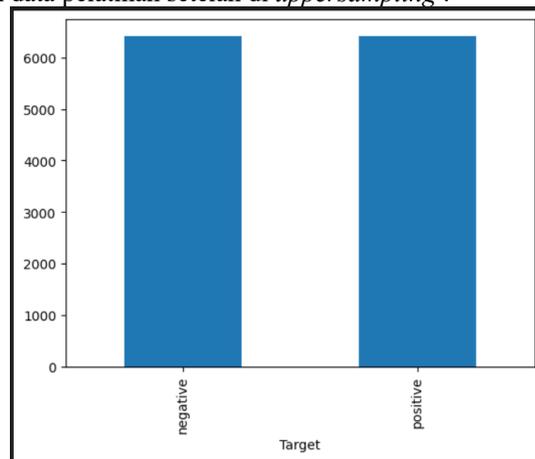
3.7 *Uppersampling*

Sebelum melakukan analisis pada data asli yang diperoleh melalui *web scraping*, model dilatih terlebih dahulu menggunakan data pelatihan (*train dataset*). Dalam proses pelatihan ini, dilakukan penyesuaian proporsi antara data positif dan negatif melalui teknik *uppersampling*, guna menjaga keseimbangan. Penyesuaian ini bertujuan untuk mencegah *overfitting* pada model, sehingga model tidak terlalu condong pada salah satu kelas, baik kelas positif maupun negatif. Dengan demikian, hasil pelatihan menjadi lebih representatif dan akurat saat diterapkan pada data asli. Pada gambar 4 merupakan data pelatihan sebelum di *uppersampling* :



Gambar 4. Sebelum *Uppersampling*

Pada gambar 5 merupakan data pelatihan setelah di *uppersampling* :



Gambar 5. Setelah *Uppersampling*

Setelah proses *uppersampling* dilakukan, *dataset negative* mengalami kenaikan data sebesar 2.980 sampel, sehingga jumlah total *dataset negative* menjadi 6.416 sampel, setara dengan jumlah *dataset positive*. Dengan demikian, jumlah keseluruhan *dataset positive* dan *negative* pada *data train* mencapai 12.832 sampel.

4. KESIMPULAN

Model SVM dilatih untuk mengklasifikasikan sentimen ulasan sebagai positif & negatif. Evaluasi model dilakukan menggunakan *confusion matrix* dengan nilai akurasi yang cukup bagus, nilai *precision* tercatat pada angka 0,89, nilai *recall* sebesar 0,90 dan nilai *f1-score* mencapai 0,89. Model ini menunjukkan performa yang baik dalam mengklasifikasikan sentimen ulasan. Dapat disimpulkan bahwa kombinasi algoritma *Support Vector Machine* (SVM) dan teknik *Term Frequency-Inverse Document Frequency* (TF-IDF) yang dijalankan menunjukkan performa yang cukup bagus dalam menjalankan analisis sentimen.

DAFTAR PUSTAKA

- [1] N. P. Rahmayanti, "Pengaruh Marketplace Dan Pembayaran Digital Terhadap Tingkat Penjualan Umkm Di Kota Banjarmasin," *Jurnal Komunikasi Bisnis dan Manajemen*, vol. 10, no. 1, 2023.

- [2] R. Valentini, P. Siwabessy, A. Herdiani, dan A. Romadhony, "Analisis Sentimen Masyarakat Terhadap Hasil Kerja Petahana Dalam Kaitan Dengan Pemilihan Presiden tahun 2019 Pada Sosial Media Twitter Menggunakan *Support Vector Machine* (SVM)," 2019.
- [3] V. Kevin, S. Que, A. Iriani, dan H. D. Purnomo, "Analisis Sentimen Transportasi Online Menggunakan *Support Vector Machine* Berbasis Particle Swarm Optimization (Online Transportation Sentiment Analysis Using *Support Vector Machine* Based on Particle Swarm Optimization)," 2020. [Daring]. Tersedia pada: www.tripadvisor.com,
- [4] I. S. K. Idris, Y. A. Mustofa, dan I. A. Salihi, "Analisis Sentimen Terhadap Penggunaan Aplikasi Shopee Menggunakan Algoritma *Support Vector Machine* (SVM)," 2023.
- [5] I. S. H. Kusuma, "Pengaruh Online Customer Review terhadap Keputusan Pembelian pada Marketplace Shopee di Kalangan Mahasiswa Kota Bandung," *International Journal Administration Business and Organization*, vol. 4, no. 2, hlm. 31–39, Agu 2023, doi: 10.61242/ijabo.23.266.
- [6] D. I. Sugiarti dan R. Iskandar, "Pengaruh Consumer Review terhadap Keputusan Pembeli Terhadap Toko Online Shopee," 2021.
- [7] E. Wolfgang, "Machine Learning for Brain Disorders," 2023. [Daring]. Tersedia pada: <http://www.springer.com/series/7657>
- [8] N. F. Putri, S. Al Faraby, dan M. Dwifabri, "Analisis Sentimen pada Produk Kecantikan dari Ulasan Female Daily Menggunakan Information Gain dan SVM Classifier," 2019.
- [9] A. Nabila, "Analisis Sentimen Ulasan Produk Toner Pada Beauty Brand 'The Body Shop' Menggunakan Metode Naïve Bayes Classifier Dan *Support Vector Machine*: Studi Kasus Di Female Daily," hlm. 1–87, 2022.
- [10] C. G. Kencana dan Y. Sibaroni, "Klasifikasi Sentiment Analysis pada Review Buku Novel Berbahasa Inggris dengan Menggunakan Metode *Support Vector Machine* (SVM)," 2019.
- [11] A. D. Auliya, S. Subanti, dan E. Zukhronah, "Implementasi Text Mining Pada Analisis Sentimen Pengguna Twitter Terhadap Marketplace di Indonesia Menggunakan Algoritma *Support Vector Machine*," 2020.
- [12] D. Deviacita, H. Sasty, dan H. Muhardi, "Implementasi Web Scraping untuk Pengambilan Data pada Situs Marketplace," vol. 7, no. 4, 2019.
- [13] D. Darwis, E. S. Pratiwi, A. Ferico, dan O. Pasaribu, "Penerapan Algoritma Svm Untuk Analisis Sentimen Pada Data Twitter Komisi Pemberantasan Korupsi Republik Indonesia," 2020.
- [14] J. A. Septian, T. M. Fahrudin, dan A. Nugroho, "Analisis Sentimen Pengguna Twitter Terhadap Polemik Persepakbolaan Indonesia Menggunakan Pembobotan TF-IDF dan K-Nearest Neighbor," 2019. [Daring]. Tersedia pada: <https://t.co/9WloaWpfD5>
- [15] S. Wulan, U. Vitandy, A. A. Supianto, dan F. Abdurrachman Bachtiar, "Analisis Sentimen Evaluasi Kinerja Dosen menggunakan *Term Frequency-Inverse Document Frequency* dan Naïve Bayes Classifier," 2019. [Daring]. Tersedia pada: <http://j-ptiik.ub.ac.id>
-