



Analisis Faktor Kepuasan Pelanggan Produk Skincare di Tiktok Shop Menggunakan *Machine Learning* dan Shap

Selma Mulky Nisa¹, Yudi Setiawan², Endina Putri Purwandari³

^{1,2,3}Program Studi Sistem Informasi, Fakultas Teknik, Universitas Bengkulu, Kota Bengkulu, Indonesia

Email: ¹selmamulkyaaa@email.com, ²ysetiawan@unib.ac.id, ³endinaputri@unib.ac.id

Abstract

Customer reviews of skincare products sold through TikTok Shop provide valuable information for evaluating customer satisfaction and understanding consumer preferences. As the volume of reviews continues to grow, manual analysis becomes increasingly inefficient, making machine learning-based sentiment analysis a more effective approach. This study applies the Random Forest algorithm to classify customer sentiment and employs Shapley Additive Explanations (SHAP) to interpret the factors influencing the model's predictions. The study utilized a dataset of 663 customer reviews collected from TikTok Shop. The analysis process consisted of text preprocessing, TF-IDF feature extraction, sentiment labeling, Random Forest model development, performance evaluation, and model interpretation using SHAP. The experimental results showed that the proposed model achieved an accuracy of 68.42%, a precision of 67.79%, a recall of 68.42%, and an F1-score of 67.29%. Of the total reviews analyzed, 341 reviews (51.43%) were classified as positive, 208 reviews (31.37%) as negative, and 114 reviews (17.19%) as neutral. SHAP analysis revealed that the terms "banget," "bagus," "bikin," "Skintific," and "kulit" were the most influential features affecting the model's predictions of customer satisfaction. These findings provide useful insights for businesses to better understand customer perceptions and support continuous improvements in product quality and service performance.

Keywords: Customer Satisfaction, Sentiment Analysis, Random Forest, SHAP, Tiktok Shop

Abstrak

Ulasan pelanggan pada produk skincare yang dipasarkan melalui TikTok Shop dapat dimanfaatkan sebagai sumber informasi untuk memahami tingkat kepuasan konsumen sekaligus mengevaluasi kualitas produk. Meningkatnya jumlah ulasan menyebabkan proses analisis secara manual tidak lagi efisien sehingga diperlukan metode yang mampu mengolah data teks secara otomatis. Dalam penelitian ini, analisis sentimen dilakukan menggunakan algoritma Random Forest, sedangkan interpretasi terhadap faktor-faktor yang memengaruhi hasil prediksi dilakukan melalui pendekatan Shapley Additive Explanations (SHAP). Penelitian memanfaatkan 663 ulasan pelanggan yang dikumpulkan dari TikTok Shop sebagai sumber data. Tahapan analisis mencakup *text preprocessing*, pembentukan fitur menggunakan TF-IDF, pelabelan sentimen, pembangunan model Random Forest, evaluasi kinerja model, serta interpretasi hasil dengan SHAP. Hasil pengujian menunjukkan bahwa model menghasilkan *accuracy* 68,42%, *precision* 67,79%, *recall* 68,42%, dan *F1-score* 67,29%. Dari keseluruhan ulasan yang dianalisis, 341 ulasan (51,43%) termasuk kategori positif, 208 ulasan (31,37%) termasuk kategori negatif, dan 114 ulasan (17,19%) termasuk kategori netral. Berdasarkan interpretasi SHAP, kata “banget”, “bagus”, “bikin”, “skintific”, dan “kulit” merupakan fitur yang memberikan pengaruh terbesar terhadap keputusan model dalam memprediksi kepuasan pelanggan. Temuan penelitian ini diharapkan dapat menjadi masukan bagi pelaku usaha untuk memahami persepsi pelanggan serta mendukung upaya peningkatan kualitas produk dan layanan.

Kata Kunci: Kepuasan Pelanggan, Analisis Sentimen, Random Forest, SHAP, Tiktok Shop.

1. PENDAHULUAN

Meningkatnya aktivitas *e-commerce* telah menyebabkan bertambahnya jumlah ulasan pelanggan di platform TikTok Shop. Ulasan-ulasan ini mencakup pengalaman pengguna serta penilaian terhadap kualitas produk, sehingga menjadikannya sumber

informasi yang penting bagi pelaku usaha (Fitria et al., 2026). Seiring bertambahnya jumlah data ulasan, analisis manual menjadi kurang efisien dan memakan waktu. Untuk mengotomatiskan pemrosesan dan analisis data teks, oleh sebab itu, penerapan *machine learning* diperlukan untuk mengolah serta menganalisis data teks secara otomatis.

Machine learning cukup efektif untuk mengklasifikasikan teks berbahasa Indonesia, termasuk mendeteksi *cyberbullying*, menurut sejumlah penelitian (Setiawan et al., 2025). Studi ini menggunakan algoritma *Forest Random* untuk klasifikasi sentimen dan pendekatan *Shapley Additive Explanations* (SHAP) untuk mengidentifikasi masing-masing faktor dalam prediksi model (Wu et al., 2025).

TikTok Shop tidak hanya berfungsi sebagai platform transaksi, tetapi juga menyediakan fitur komentar dan ulasan yang dapat dimanfaatkan pengguna untuk berbagi pengalaman mereka terkait produk. Ulasan-ulasan ini mencakup penilaian konsumen terhadap kualitas, manfaat, dan kinerja suatu produk, sehingga menjadikannya sumber informasi yang berharga baik bagi calon pembeli maupun pemilik usaha. Namun, seiring dengan meningkatnya jumlah ulasan yang ditulis dalam bahasa bebas, analisis manual menjadi kurang efektif dan lebih memakan waktu. Oleh karena itu, penelitian ini menggunakan analisis sentimen untuk mengkategorikan umpan balik konsumen secara sistematis sehingga dapat digunakan untuk mengevaluasi produk (Sari et al., 2025).

Sari et al. (2025) menggunakan algoritma *Random Forest* dan *Naïve Bayes* untuk mengklasifikasikan sentimen pada ulasan produk perawatan kulit di TikTok Shop. Model *Random Forest* mencapai akurasi sebesar 82%, sedangkan model *Naïve Bayes* mencapai akurasi sebesar 62%. Dengan demikian, model *Random Forest* lebih baik dalam memproses data ulasan daripada *Naïve Bayes*, yang menghasilkan klasifikasi sentimen yang lebih baik.

Dalam analisis sentimen ulasan aplikasi TikTok, Tangke et al. (2024) meneliti kinerja *Random Forest* mencapai akurasi sebesar 93%. Hasilnya menunjukkan bahwa *Super Vector Machine* (SVM) memiliki akurasi lebih rendah. Kondisi tersebut membuat *Random Forest* lebih baik daripada metode lain untuk mengklasifikasikan sentimen. Keunggulan ini menyoroti kemampuan *Random Forest* dalam mengakomodasi keragaman karakteristik data ulasan sekaligus menjaga stabilitas model selama proses klasifikasi, sehingga mengurangi risiko *overfitting*. Sehingga, metode *Random Forest* dianggap tepat untuk analisis sentimen yang melibatkan data teks dalam jumlah besar (Tangke et al., 2024).

Dalam penelitian, Winanta et al. (2026) membangun model analisis sentimen untuk ulasan *e-commerce* dengan memanfaatkan algoritma *Random Forest* yang dipadukan dengan pendekatan *Explainable Artificial Intelligence* (XAI) melalui metode LIME, SHAP, dan ELI5. Hasil penelitian menunjukkan bahwa *Random Forest* efektif dalam mengklasifikasikan sentimen. Selain itu, teknik SHAP digunakan untuk menginterpretasikan seberapa besar kontribusi setiap fitur terhadap hasil prediksi, sehingga proses pengambilan keputusan model menjadi lebih transparan dan memudahkan identifikasi faktor-faktor yang paling berpengaruh terhadap prediksi sentimen.

Meskipun metode *Deep Learning* dan Transformer sering digunakan dalam analisis sentimen karena kinerjanya yang unggul, implementasinya umumnya memerlukan dataset yang lebih besar dan sumber daya komputasi yang lebih banyak. Dengan dataset sebanyak 663 ulasan pelanggan, penelitian ini menggunakan *Random Forest* karena metode ini berkinerja baik pada dataset kecil, memiliki proses pelatihan yang lebih cepat, dan dapat dipadukan dengan SHAP untuk memberikan interpretasi terhadap hasil prediksi. Oleh karena itu, pemilihan metode ini didasarkan pada keseimbangan antara kinerja klasifikasi, efisiensi komputasi, dan kemampuan untuk menjelaskan hasil prediksi sesuai dengan tujuan penelitian.

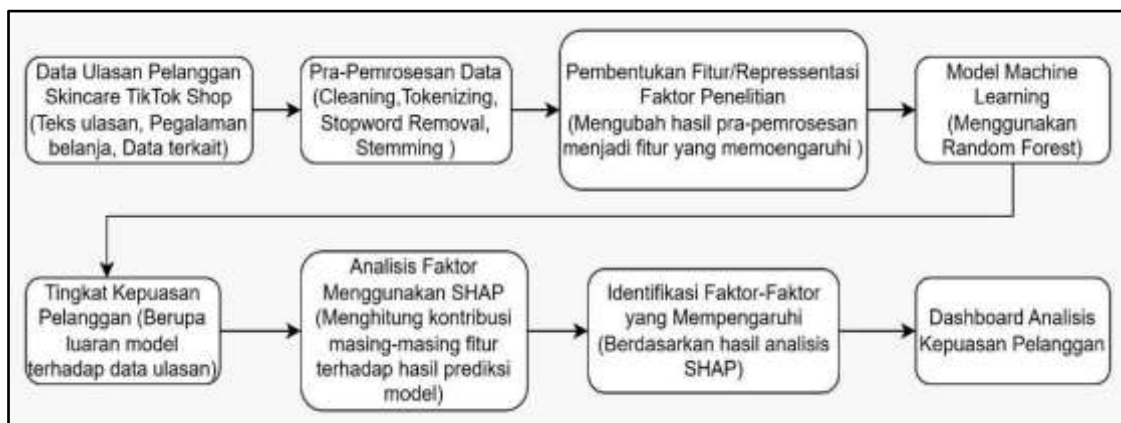
Beberapa penelitian telah menunjukkan bahwa *Random Forest* memiliki kinerja yang baik dalam klasifikasi sentimen. Namun, SHAP meningkatkan interpretabilitas model dengan menjelaskan kontribusi setiap faktor terhadap hasil prediksi (Wu et al, 2025). Di sisi lain, analisis sentimen terhadap ulasan *e-commerce* dapat membantu mengidentifikasi perspektif pelanggan dan tingkat kepuasan mereka (Chandra, 2024). Namun, penelitian yang menggabungkan *Random Forest* dan SHAP untuk mengidentifikasi karakteristik yang memengaruhi kepuasan pelanggan berdasarkan penilaian produk *skincare* di TikTok Shop masih terbatas. Oleh karena itu, penelitian ini menggunakan kedua metode tersebut untuk menghasilkan klasifikasi sentimen yang lebih mudah diinterpretasikan sekaligus mengidentifikasi faktor-faktor yang memengaruhi kepuasan pelanggan.

Penelitian ini bertujuan untuk: (1) membangun model *machine learning* menggunakan algoritma *Random Forest* untuk memprediksi tingkat kepuasan pelanggan produk *skincare* di TikTok Shop berdasarkan data ulasan pelanggan; (2) menganalisis dan mengidentifikasi faktor-faktor yang memengaruhi tingkat kepuasan pelanggan produk *skincare* di TikTok Shop menggunakan metode SHAP; (3) menjelaskan kontribusi masing-masing faktor terhadap hasil prediksi model *Random Forest* menggunakan metode SHAP sebagai pendekatan XAI; dan (4) menyajikan hasil analisis dalam bentuk *dashboard* yang memuat visualisasi hasil prediksi, faktor-faktor dominan yang memengaruhi kepuasan pelanggan, serta hasil interpretasi SHAP untuk mempermudah pemahaman hasil penelitian. Hasil penelitian ini diharapkan dapat membantu pelaku bisnis menemukan faktor-faktor yang memengaruhi kepuasan pelanggan dan membantu dalam pengembangan produk, peningkatan layanan, dan strategi pemasaran yang lebih baik.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Penelitian ini dilakukan dalam beberapa langkah yang saling terkait untuk mengembangkan model analisis sentimen yang mengidentifikasi faktor-faktor yang memengaruhi kepuasan pelanggan terhadap *skincare* TikTok Shop. Tahapan ini dimulai dengan pengumpulan data, melalui pemrosesan teks, ekstraksi TF-IDF dari fitur, klasifikasi sentimen, dan pembuatan model *Random Forest*. Akhir dari tahapan ini adalah evaluasi kinerja modifikasi. Tahapan penelitian disusun secara sistematis mengacu pada alur pengembangan analisis sentimen berbasis *machine learning* yang telah banyak diterapkan pada penelitian sebelumnya (Angkosos et al., 2024).



Gambar 1. Tahapan Penelitian

Gambar 1 menunjukkan tahapan penelitian yang digunakan. Penelitian dimulai dari pengumpulan data ulasan skincare TikTok Shop, diikuti dengan pra-pemrosesan (*cleaning, tokenizing, stopword removal, dan stemming*) untuk meningkatkan kualitas data. Data yang sudah dikumpulkan kemudian dikonversi ke nilai numerik menggunakan TF-IDF sebelum dikategorikan menggunakan metode *Random Forest*. Hasil klasifikasi kemudian dianalisis menggunakan SHAP untuk menentukan kontribusi setiap fitur terhadap prediksi model. Data ini digunakan untuk mengidentifikasi faktor-faktor yang memengaruhi kepuasan pelanggan dan disajikan dalam sebuah *dashboard* untuk mempermudah pemahaman hasil analisis.

2.2 Dataset Penelitian

Penelitian ini menganalisis 663 ulasan pengguna terhadap produk skincare yang dikumpulkan dari TikTok Shop. Data yang dikumpulkan mencakup *brand*, nama produk, dan ulasan pengguna. Sebelum digunakan untuk membuat model klasifikasi sentimen, data diproses melalui *text preprocessing*, ekstraksi fitur TF-IDF, dan pelabelan sentimen. Hasil dari proses tersebut menjadi masukan bagi algoritma *Random Forest*, untuk menganalisis faktor-faktor yang memengaruhi kepuasan pelanggan. Tabel 1 menunjukkan contoh karakteristik data penelitian.

Tabel 1. Karakteristik Data

Karakteristik	Keterangan
Sumber Data	TikTok Shop
Objek Penelitian	Ulasan Pelanggan Produk Skincare
Jumlah Data	663 Ulasan
Format Data	Data Teks
Variabel Utama	Brand, Produk, Ulasan
Jenis Data	Data Sekunder

2.3 Text PreProcessing

Sebelum data digunakan untuk membangun model klasifikasi, dilakukan tahapan *text preprocessing* guna mempersiapkan kualitas data teks agar lebih sesuai untuk proses analisis. Tahapan ini mencakup penghapusan informasi yang tidak diperlukan, penyeragaman bentuk kata, serta penyusunan data menjadi lebih konsisten sehingga proses klasifikasi dapat dilakukan secara lebih optimal. Rangkaian *text preprocessing* yang diterapkan terdiri atas *cleaning, tokenizing, stopword removal, dan stemming*.

- Cleaning* dilakukan untuk mempersiapkan data teks agar lebih siap dianalisis dengan menghilangkan berbagai karakter yang tidak memiliki nilai informasi, seperti simbol, angka, tanda baca, dan tautan. Selain itu, seluruh huruf diubah menjadi huruf kecil untuk menjaga konsistensi penulisan.
- Tokenizing* dilakukan untuk memisahkan kalimat ulasan menjadi kumpulan kata sehingga setiap kata dapat diproses secara terpisah pada tahap berikutnya.
- Stopword Removal* untuk menghilangkan kata-kata umum yang tidak relevan untuk analisis sentimen, sehingga hanya kata-kata yang mengandung informasi yang relevan yang tersisa.
- Stemming* merupakan tahapan yang menyederhanakan setiap kata berimbuhan menjadi bentuk dasarnya.

2.4 TF-IDF

Pendekatan TF-IDF memberikan bobot pada setiap kata untuk mengubah data ulasan menjadi representasi numerik, dan hasil pra-pemrosesan teks dapat digunakan untuk melatih model *machine learning* (Munaldi, 2025). Bobot ini didasarkan pada

berapa kali suatu kata muncul dalam satu ulasan atau di seluruh data ulasan. Menurut Munaldi (2025), kata-kata yang sering muncul di banyak ulasan tetapi tidak di ulasan lain harus diberi nilai yang lebih tinggi karena lebih mencerminkan isi ulasan tersebut. Selain itu, kata-kata yang sering muncul di banyak ulasan harus diberi nilai yang lebih rendah karena tidak sesuai untuk membedakan kualitas ulasan.

Representasi numerik yang diperoleh dari pembobotan TF-IDF disusun ke dalam matriks fitur dan dimasukkan ke dalam algoritma *Random Forest* untuk membuat model klasifikasi sentimen terhadap ulasan pelanggan (Munaldi, 2025).

2.5 Labelling

Setelah semua tahapan *preprocessing* selesai, model *Indonesian RoBERTa Base Sentiment Classifier* dari pustaka *Transformers* (*w11wo/indonesian-roberta-base-sentiment-classifier*) secara otomatis memberikan label sentimen pada setiap ulasan. Dibandingkan dengan metode manual, model yang telah dilatih sebelumnya ini dapat memberikan pelabelan yang lebih konsisten dan efektif karena dirancang khusus untuk klasifikasi sentimen pada teks berbahasa Indonesia (Pratama et al., 2025).

Setiap ulasan dikategorikan ke dalam tiga kategori sentimen seperti, positif, negatif, atau netral. Model *Random Forest* kemudian dilatih dan diuji menggunakan label tersebut untuk menentukan bagaimana katakarakteristik teks dan kelas sentimen saling berhubungan. Penggunaan *Indonesian RoBERTa Base Sentiment Classifier* juga mendukung anotasi otomatis pada data berbahasa Indonesia secara lebih efisien (Nainggolan & Handoko, 2025).

Validasi manual dilakukan terhadap 50 ulasan yang dipilih secara acak dari total 663 data ulasan guna menjamin keakuratan pelabelan otomatis. Validasi dilakukan dengan membandingkan hasil pelabelan *Indonesian RoBERTa Base Sentiment Classifier* dengan penilaian manual berdasarkan makna keseluruhan ulasan. Dalam proses ini, ulasan diklasifikasikan sebagai negatif jika berisi keluhan atau ketidakpuasan, netral jika hanya berisi fakta atau pertanyaan tanpa kecenderungan sentimen yang jelas, dan positif jika menunjukkan kepuasan, rasa terima kasih, atau rekomendasi terhadap produk.

Berdasarkan hasil validasi, 46 dari 50 ulasan memiliki label yang sesuai dengan penilaian manual dan pelabelan otomatis, sedangkan 4 ulasan menunjukkan perbedaan. Temuan pelabelan otomatis cukup konsisten untuk digunakan sebagai acuan dalam proses klasifikasi algoritma *Random Forest*, sebagaimana ditunjukkan oleh tingkat kesesuaian sebesar 92%. Hasil Pelabelan terdapat pada Tabel 2.

Tabel 2. Hasil Validasi Pelabelan Sentimen

Keterangan	Jumlah
Jumlah Sampel Validasi	50
Label sesuai	46
Label Tidak Sesuai	4
Tingkat Kesesuaian	92%

2.6 Random Forest

Pemodelan dilakukan setelah ulasan dikategorikan berdasarkan sentimennya dan diubah menjadi fitur numerik menggunakan TF-IDF. Data kemudian dibagi menjadi dua bagian: 80% data latih dan 20% data uji. Berdasarkan data pelatihan, model klasifikasi dibangun menggunakan metode *Random Forest*, dan data uji digunakan untuk mengevaluasi kemampuan model untuk mengkategorikan sentimen ulasan produk skincare TikTok Shop.

SMOTE, *undersampling*, dan *class weighting* adalah contoh strategi penyeimbangan data yang tidak digunakan selama fase pelatihan guna mempertahankan distribusi kelas pada dataset asli. Metode ini dikembangkan agar model *Random Forest* dapat mengidentifikasi pola sentimen berdasarkan distribusi data yang dikumpulkan. Terdapat 114 ulasan netral (17,19%), 208 ulasan negatif (31,37%), dan 341 ulasan positif (51,43%) dalam dataset. Ketidakseimbangan tersebut dapat memengaruhi kemampuan model dalam mengenali kelas minoritas, terutama kelas netral, sehingga performa klasifikasinya cenderung lebih rendah, sebagaimana terlihat pada *Confusion Matrix*.

Sebelum menetapkan model akhir, dilakukan pengujian terhadap konfigurasi parameter *Random Forest* untuk mengevaluasi pengaruhnya terhadap kinerja model. Parameter yang diuji meliputi jumlah pohon keputusan (*n_estimators*), kedalaman maksimum pohon (*max_depth*), jumlah minimum sampel untuk melakukan pemisahan (*min_samples_split*), serta jumlah minimum sampel pada setiap daun (*min_samples_leaf*). Pengujian dilakukan menggunakan konfigurasi *n_estimators* = 500, *max_depth* = 20, *min_samples_split* = 5, *min_samples_leaf* = 2, dan *random_state* = 42.

Hasil pengujian menunjukkan bahwa konfigurasi ini memperoleh *accuracy* 68,42%, *precision* 67,79%, *recall* 68,42%, dan *F1-Score* 67,29%. Oleh karena itu, konfigurasi dipilih sebagai model terakhir karena lebih baik dalam mengklasifikasikan dataset ini. Selanjutnya, konfigurasi akan digunakan untuk evaluasi dan analisis model dengan SHAP.

Algoritma *Random Forest* dipilih karena mampu menghasilkan prediksi yang stabil, memiliki kemampuan generalisasi yang baik, dan mengurangi potensi *overfitting*. Pemilihan metode ini didukung oleh penelitian Sari et al. (2025) serta hasil pengujian parameter pada penelitian ini yang menunjukkan bahwa konfigurasi awal memberikan performa terbaik. Selanjutnya, model dievaluasi menggunakan data uji berdasarkan nilai *accuracy*, *precision*, *recall*, dan *F1-score*.

3. HASIL DAN PEMBAHASAN

3.1 Klasifikasi Sentimen

Setelah pelabelan sentimen selesai dilakukan, setiap ulasan memperoleh salah satu dari tiga kategori, yaitu positif, negatif, atau netral. Ringkasan distribusi hasil klasifikasi terhadap 663 ulasan pelanggan disajikan kedalam bentuk tabel.

Tabel 3. Hasil Klasifikasi Sentimen

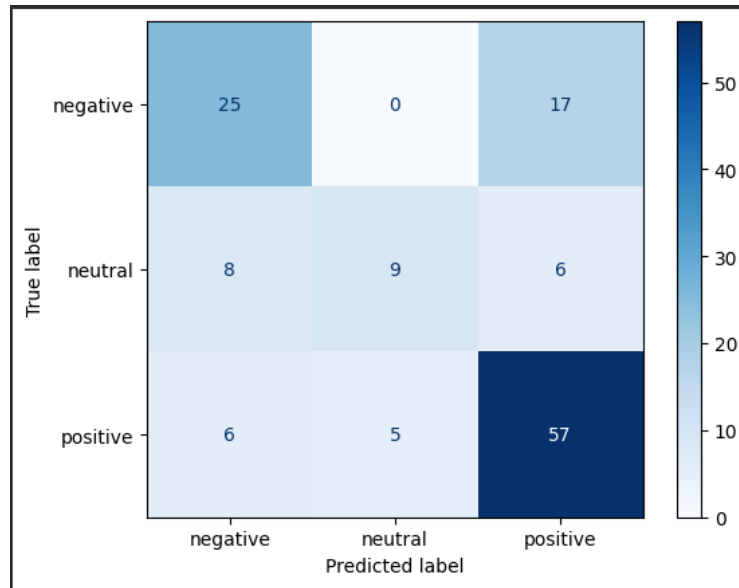
Sentimen	Jumlah	Persentase%
Positif	341	51,43%
Negatif	208	31,37%
Netral	114	17,19%
Total	663	100%

Berdasarkan data tabel 2 kategori sentimen positif mendominasi dengan jumlah 341 ulasan (51,43%), diikuti sentimen negatif sebanyak 208 ulasan (31,37%) dan sentimen netral sebanyak 114 ulasan (17,19%). Dominasi sentimen positif tersebut mengindikasikan bahwa persepsi pelanggan terhadap produk *skincare* di TikTok Shop cenderung baik.

3.2 Evaluasi Model

Untuk mengevaluasi performa model random forest dengan menggunakan *Confusion Matrix* untuk membandingkan hasil prediksi terhadap label sentimen pada data uji. Dengan menggunakan *Confusion Matrix* ini, nilai *accuracy*, *precision*, *recall*, dan

F1-score dihitung sebagai indikator performa model. Hasil evaluasi ini kemudian digunakan sebagai dasar interpretasi metode SHAP.



Gambar 2. *Confussion Matrix*

Gambar 2 menyajikan hasil evaluasi model *Random Forest* berdasarkan *confusion matrix*. Berdasarkan hasil pengujian, model mampu mengklasifikasikan dengan benar sebanyak 57 ulasan bersentimen positif, 25 ulasan bersentimen negatif, dan 9 ulasan bersentimen netral. Namun demikian, beberapa kesalahan prediksi masih ditemukan, terutama pada kelas netral yang lebih sering dipetakan ke kategori positif maupun negatif. Kondisi tersebut menunjukkan bahwa model lebih mudah mengenali karakteristik ulasan dengan sentimen positif dibandingkan ulasan yang bersifat netral. Secara keseluruhan, pola yang terlihat pada *confusion matrix* sejalan dengan hasil evaluasi yang menunjukkan *accuracy* sebesar 68,42%, sehingga algoritma *Random Forest* mampu menghasilkan performa klasifikasi yang cukup memadai untuk menganalisis sentimen ulasan pelanggan terhadap produk *skincare* di TikTok Shop.

Kesalahan klasifikasi pada kelas netral diduga terjadi karena karakteristik ulasan pelanggan di TikTok Shop cenderung menggunakan bahasa yang singkat, tidak langsung, dan sering kali hanya berisi pertanyaan, pernyataan informatif, atau pengalaman yang tidak secara tegas menunjukkan kecenderungan sentimen tertentu. Kondisi tersebut menyebabkan batas antara sentimen netral dengan sentimen positif maupun negatif menjadi kurang jelas, sehingga model mengalami kesulitan dalam membedakan ketiga kelas tersebut secara konsisten.

Selain itu, beberapa ulasan juga mengandung ungkapan yang bersifat ambigu atau bergantung pada konteks penggunaan produk, sehingga makna sentimen tidak dapat ditentukan hanya berdasarkan keberadaan kata-kata tertentu. Karena penelitian ini menggunakan representasi fitur TF-IDF yang berfokus pada frekuensi kemunculan kata, hubungan antar kata dan konteks kalimat belum dapat dipahami secara menyeluruh. Akibatnya, sebagian ulasan netral diprediksi sebagai sentimen positif maupun negatif, sebagaimana ditunjukkan pada hasil *Confusion Matrix*.

Hasil evaluasi menunjukkan bahwa model *Random Forest* memperoleh *accuracy* 68,42%, *precision* 67,79%, *recall* 68,42%, dan *F1-score* 67,29%, yang mengindikasikan kemampuan klasifikasi sentimen tergolong cukup baik. Namun, model masih mengalami kesalahan prediksi, terutama pada kelas sentimen netral karena karakteristiknya lebih sulit dibedakan dibandingkan sentimen positif dan negatif. Untuk meningkatkan kinerja

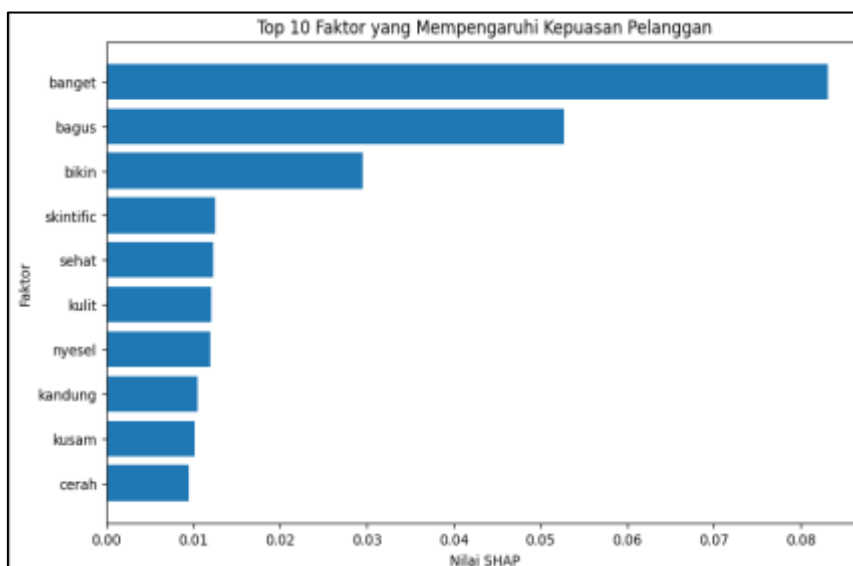
model, penelitian selanjutnya dapat mempertimbangkan penambahan data pelatihan, penerapan teknik penyeimbangan kelas, atau penggunaan metode representasi teks yang lebih mampu memahami konteks kalimat.

Nilai *accuracy* sebesar 68,42% menunjukkan bahwa *Random Forest* mampu mengklasifikasikan sebagian besar ulasan dengan cukup baik, meskipun masih terdapat beberapa prediksi yang tidak sesuai dengan label sebenarnya. Hasil tersebut dipengaruhi oleh penggunaan TF-IDF sebagai representasi fitur, yang efektif memberikan bobot pada kata-kata penting berdasarkan frekuensi kemunculannya, tetapi belum mampu merepresentasikan hubungan antarkata maupun konteks kalimat secara menyeluruh. Kondisi ini menyebabkan ulasan yang bersifat implisit, ambigu, atau bergantung pada konteks masih berpotensi mengalami kesalahan klasifikasi.

Selain keterbatasan TF-IDF, ketidakseimbangan distribusi kelas turut memengaruhi kinerja model. Dataset terdiri atas 341 ulasan positif (51,43%), 208 ulasan negatif (31,37%), dan 114 ulasan netral (17,19%), sehingga kelas netral memiliki jumlah data paling sedikit. Kondisi ini membatasi kemampuan model dalam mempelajari karakteristik sentimen netral, yang berdampak pada rendahnya akurasi klasifikasi pada kelas tersebut. Temuan ini didukung oleh *Confusion Matrix*, yang menunjukkan bahwa sebagian ulasan netral masih diklasifikasikan sebagai sentimen positif atau negatif.

3.3 Analisis Faktor Menggunakan SHAP

Setelah evaluasi model *Forest Random*, pendekatan SHAP digunakan untuk menilai prediksi model. *SHAP Feature Importance Bar Plot* menunjukkan faktor-faktor yang paling penting, sedangkan *Beeswarm Plot*, menunjukkan distribusi dan arah pengaruh masing-masing faktor terhadap prediksi. Penelitian ini menghasilkan klasifikasi sentimen dan faktor-faktor yang memengaruhi kepuasan pelanggan skincare di TikTok Shop dengan menggunakan pendekatan SHAP untuk menginterpretasikan model.



Gambar 3. SHAP Feature Importance Bar Plot

Gambar 3 menunjukkan hasil visualisasi yang menunjukkan besarnya kontribusi setiap fitur terhadap prediksi sentimen. Hasil analisis memperlihatkan bahwa kata "banget", "bagus", dan "bikin" memiliki nilai SHAP tertinggi sehingga memberikan pengaruh terbesar terhadap prediksi model. Fitur lain, seperti "skintific", "sehat", "kulit", "nyesel", "kandung", "kusam", dan "cerah", juga berkontribusi, meskipun dengan tingkat pengaruh yang lebih rendah.

Dominasi kata "banget" dan "bagus" menunjukkan bahwa pelanggan TikTok Shop cenderung menggunakan bahasa yang sederhana dan ekspresif dalam menulis ulasan produk skincare. Kata "bagus" merepresentasikan kepuasan terhadap kualitas produk, sedangkan "banget" berperan sebagai penegas untuk memperkuat sentimen positif. Temuan ini mengindikasikan bahwa pelanggan lebih sering menyampaikan kepuasan melalui ungkapan sehari-hari yang mudah dipahami.

Hasilnya memberikan informasi yang memudahkan pelaku usaha skincare di TikTok Shop. Informasi ini memudahkan untuk melihat faktor-faktor yang paling berpengaruh untuk menilai kualitas produk, membuat strategi periklanan yang lebih sesuai dengan persepsi pelanggan, dan memungkinkan evaluasi produk serta layanan yang lebih tepat sasaran.

Interpretasi menggunakan SHAP menunjukkan bahwa prediksi *Random Forest* dipengaruhi oleh kata-kata yang mencerminkan pengalaman dan penilaian pelanggan terhadap produk. Temuan ini menunjukkan bahwa model tersebut tidak hanya mampu mengklasifikasikan sentimen, tetapi juga mengidentifikasi faktor-faktor yang memiliki pengaruh terbesar terhadap prediksi, yang meningkatkan transparansi model dan memudahkan pengambilan keputusan.

3.4 Dashboard Analisis Kepuasan Pelanggan

Hasil penelitian disajikan dalam *dashboard* interaktif yang menampilkan distribusi sentimen, kinerja *Random Forest* (*accuracy*, *precision*, *recall*, dan *F1-score*), serta fitur yang paling berkontribusi terhadap prediksi berdasarkan interpretasi SHAP. Tujuannya untuk memudahkan pengguna memahami hasil analisis sekaligus memberikan informasi bagi pelaku usaha untuk mengevaluasi persepsi pelanggan serta mendukung pengembangan produk dan layanan.



Gambar 4. *Dashboard*

Gambar 4 menampilkan *dashboard* yang mengintegrasikan hasil penelitian sehingga analisis lebih mudah dipahami. *Dashboard* menyajikan jumlah ulasan yang dianalisis, kinerja model *Random Forest*, distribusi sentimen pelanggan, serta faktor-faktor yang memengaruhi kepuasan pelanggan berdasarkan interpretasi SHAP. Penyajian ini memudahkan pengguna memahami hasil penelitian tanpa perlu meninjau setiap tahapan analisis secara terpisah.

Dashboard menunjukkan bahwa penelitian ini menggunakan 663 ulasan pelanggan produk skincare di TikTok Shop. Hasil evaluasi *Random Forest* memperoleh *accuracy* 68,42%, *precision* 67,79%, *recall* 68,42%, dan *F1-score* 67,29%. Evaluasi model menunjukkan bahwa *Random Forest* mampu menghasilkan klasifikasi sentimen dengan kinerja yang baik. Hasil tersebut kemudian digunakan sebagai dasar untuk menganalisis kepuasan pelanggan.

Distribusi sentimen menunjukkan bahwa ulasan positif mendominasi sebesar 51,43%, diikuti sentimen negatif 31,37% dan netral 17,19%. Hasil tersebut mengindikasikan bahwa pelanggan cenderung memberikan penilaian positif terhadap produk skincare di TikTok Shop, sedangkan ulasan negatif dan netral tetap memberikan masukan yang berguna untuk pengembangan produk serta peningkatan kualitas layanan.

Hasil interpretasi SHAP menunjukkan bahwa kata "banget", "bagus", "bikin", "skintific", dan "kulit" merupakan fitur yang paling berkontribusi terhadap prediksi sentimen. Temuan tersebut mengindikasikan bahwa pengalaman penggunaan, kualitas produk, dan kesesuaiannya dengan kondisi kulit menjadi faktor yang berkaitan dengan kepuasan pelanggan. Selain itu, *dashboard* yang dikembangkan menyajikan hasil analisis secara ringkas dan terstruktur sehingga memudahkan pelaku usaha dalam mengevaluasi kepuasan pelanggan berdasarkan ulasan.

Hasil penelitian ini sejalan dengan penelitian Winanta et al. (2026), yang mengklasifikasikan sentimen *e-commerce* dengan menggunakan metode XAI seperti LIME, SHAP, dan ELI5 pada model *Random Forest*. Menurut penelitian ini, SHAP dapat menjelaskan peran setiap fitur dalam prediksi, meningkatkan transparansi dan interpretasi model. Kesesuaian hasil ini mendukung kinerja SHAP dalam mengidentifikasi faktor-faktor yang memengaruhi kepuasan pelanggan berdasarkan data ulasan.

4. KESIMPULAN

Penelitian ini menunjukkan bahwa *Random Forest* mampu mengidentifikasi sentimen dalam ulasan produk skincare di TikTok Shop dengan *accuracy* 68,42%, *precision* 67,79%, *recall* 68,42%, dan *F1-score* 67,29%. Interpretasi SHAP mengidentifikasi kata-kata "banget," "bagus," "bikin," "skintific," dan "kulit" sebagai fitur yang paling berkontribusi terhadap prediksi sentimen, sementara hasil analisis menunjukkan dominasi sentimen positif. Lebih lanjut, *dashboard* yang dirancang memudahkan interpretasi hasil analisis dengan menyajikannya secara terintegrasi. Untuk penelitian lebih lanjut, disarankan untuk menggunakan sumber data yang lebih beragam dan membandingkan berbagai metode pembelajaran mesin untuk meningkatkan kinerja model, karena penelitian saat ini terbatas pada evaluasi TikTok Shop dan menggunakan satu algoritma saja.

REFERENCES

- Chandra, W. (2024). *Kepuasan Pelanggan Di Tiktok Shop*. *Jurnal Manajerial dan Kewirausahaan* 06(02), 552–558.
- Cucun Very Angkoso, Mochamad Adrian Nuradha Thrisna, Budi Dwi Satoto, A. K. (2024). *Optimasi Klasifikasi Sentimen Menggunakan*. *Jurnal Edukasi dan Penelitian Informatika*, 10(3), 389–400.
- Dewi, T., & Afrillia, Y. (2025). *Sentiment Analysis of Customer Satisfaction Towards Shopee and Lazada E-commerce Platform Using the Random Forest Algorithm Classifier*. *International Journal of Engineering, Science and Information Technology*, 5(1), 229–235. <https://doi.org/10.52088/ijesty.v5i1.692>

- Fitria, N., Nurkasi, S. L., & Aulia, N. D. (2026). *Pengaruh Kualitas Sistem TikTok Shop terhadap Kepuasan Pelanggan Dalam Pembelian Produk Fashion Online*. 5(1), 45–51.
- Harmandini, K. P., & L, K. M. (2024). *Analysis of TF-IDF and TF-RF Feature Extraction on Product Review Sentiment*. *Sinkron: Jurnal dan Penelitian Teknologi Informatika*, 8(2), 929–937.
- Iffadah, A. S., & Prasetya, D. A. (2025). *Shapley Additive Explanations Interpretation of the XGBoost Model in Predicting Air Quality in Jakarta*. *Jurnal Riset Informatika*, 7(3). <https://doi.org/10.34288/jri.v7i3.366>
- Komputer, F. I., & Pamulang, U. (2025). *Penerapan Algoritma Random Forest Berbasis TF – IDF untuk Meningkatkan untuk Akurasi Deteksi Sentimen Ulasan Produk E- Commerce*. *Riau Jurnal Informatika*, 4(3), 454–459.
- Lestari, V. B., & Hutagalung, C. A. (n.d.). *Evaluation of TF-IDF Extraction Techniques in Sentiment Analysis of Indonesian-Language Marketplaces Using SVM, Logistic Regression, and Naive Bayes*. *Journal of Computer Science and Applications*, 021, 36–44. <https://doi.org/10.21009/j-koma.v8i1.05>
- Nur, M. A., Umar, N., Feng, Z., & Gani, H. (2025). *Evaluation Of Indobert And Roberta : Performance Of Indonesian Language Transformer Models In Sentiment*. *JIKO: Jurnal Informatika dan Komputer*, 8(0173), 121–127. <https://doi.org/10.33387/jiko.v8i2.9988>
- Puspanagara, A. L. (2025). *Penerapan Explainable AI untuk Prediksi Performa Akademik Mahasiswa Menggunakan Random Forest dan SHAP*. *Infoman's: Jurnal Ilmu-ilmu Informatika dan Manajemen*, 19(1), 1–7.
- Saputra, F. (2025). *Prediction Customer Loyalty Using Random Forest Algorithm on Shopee Reviews*. *Jurnal Universitas Bina Sarana Informatika*, 27(1), 11–20. <https://doi.org/10.31294/p.v27i1.7940>
- Sari, P. W., Kadafi, A. R., Bina, U., & Informatika, S. (2025). *Perbandingan Algoritma Random Forest Dan Naïve Bayes Dalam Menganalisis Sentimen Ulasan Pada Produk Skincare Lokal Di Media*. *JITET (Jurnal Informatika dan Teknik Elektro Terapan)* 13(3), 1946–1957. <http://dx.doi.org/10.23960/jitet.v13i3S1.8150>
- Setiawan, Y., Purwandari, E. P., Wijanarko, A., Panca, Y. P., & Utama, F. P. (2025). *Detection of Bahasa cyberbullying speech using large-scale N-gram machine learning models with increased document-terms probability*. *Bol. Soc. Paran. Mat.*, 43(4), 1–14. <https://doi.org/10.5269/bspm.78695>
- Winanta, H. P., Hana, M. Y., & Hasan, F. N. (2026). *Comparative Analysis of Explainable AI Methods LIME, SHAP, and ELI5 on Random Forest Based Indonesian E-Commerce Sentiment Classification*. *Jurnal Teknik Informatika (JUTIF)*, 7(2), 1778–1799. <https://doi.org/10.52436/1.jutif.2026.7.2.5642>
- Wu, Q., Xia, C., & Tian, S. (n.d.). *AI-Driven Sentiment Analytics : Unlocking Business Value in the E-Commerce Landscape*.
- Wulandary, E., Mas, M., & Ashoer, M. (2023). *Pengaruh Kualitas Layanan Era Digitalisasi Media Pemasaran Online E-Commerce dengan Menggunakan E-Service Quality Untuk Mengetahui Kepuasan Pelanggan*. *Center of Economic Student Journal*, 6(1), 44–59. <https://doi.org/10.56750/csej.v6i1.568>