

Analisis Sentimen Ulasan Produk Kecantikan di Tokopedia Menggunakan IndoBERT

Ayumi¹, Angela Fayola², Rosita Darianty³, Stephanie⁴

¹ Sistem Informasi, Universitas Pelita Harapan, 03081220019@student.uph.edu

² Sistem Informasi, Universitas Pelita Harapan, 03081220023@student.uph.edu

³ Sistem Informasi, Universitas Pelita Harapan, 03081220014@student.uph.edu

⁴ Sistem Informasi, Universitas Pelita Harapan, 03081220016@student.uph.edu

Article Info:

Article history:

Received Date: 20/04/2025

Accepted Date: 28/04/2025

Published Date: 31/05/2025

Keywords:

IndoBERT

Sentiment Analysis

Beauty Products

ABSTRACT

The rapid growth of information and communication technology has significantly influenced consumer behaviour, especially in online shopping. In Indonesia, e-commerce has transformed how customers' purchase beauty products, with platforms like Tokopedia playing an important role in facilitating these transactions. However, the increasing volume of customer reviews poses challenges for businesses to analyze consumer sentiment effectively. This research aims to analyze customer reviews of Skintific moisturizer products using Natural Language Processing (NLP) with the IndoBERT model, which is specifically trained for the Indonesian language. Using a dataset of 1,394 reviews, the research adopts a mixed-method approach, utilizing qualitative sentiment analysis for customer reviews and quantitative analysis for star ratings. The sentiment analysis process involves web scraping (ParseHub), text preprocessing, and sentiment classification with the findings show that 72,7% of reviews were positive, 17,4% negative, and 9,9% neutral. Additionally, star ratings strongly correlated with sentiment labels, where 5-star reviews were predominantly positive (85,86%) and 1-star reviews mostly negative (85,96%). The study concludes that IndoBERT effectively classifies sentiments in Indonesian-language reviews, providing valuable insights for businesses to enhance product strategies and customer satisfaction.

This is a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

Corresponding Author:

Rosita Darianty

Universitas Pelita Harapan

03081220014@student.uph.edu

1. PENDAHULUAN

Perkembangan teknologi informasi dan komunikasi telah membawa dampak signifikan terhadap pola konsumsi masyarakat, khususnya dalam hal berbelanja. Salah satu bentuk transformasi yang paling menonjol adalah munculnya *electronic commerce* (*e-commerce*) seperti Tokopedia. *E-commerce* merupakan aktivitas pembelian, penjualan, pemasaran produk dan layanan, serta informasi melalui jaringan internet dengan memanfaatkan sistem digital [1]. *E-commerce* kini menjadi bagian dari gaya hidup masyarakat modern, termasuk di Indonesia, dalam melakukan transaksi jual beli secara digital [2]. Seiring perkembangannya, *e-commerce* tidak hanya digunakan untuk memenuhi kebutuhan pokok, tetapi juga mencakup produk-produk lainnya seperti kosmetik dan perawatan diri. Industri kecantikan sendiri telah mengalami pertumbuhan yang pesat, dengan *e-commerce* yang memainkan peran utama dalam ekspansi ini. Hal ini terbukti dengan penjualan produk kecantikan melalui platform digital yang mengalami peningkatan sekitar empat kali lipat pada kurun waktu 2015 – 2022 [3].

Dalam konteks pembelian produk kecantikan melalui *e-commerce*, kepuasan pelanggan menjadi faktor krusial yang memengaruhi tingkat kepercayaan, serta loyalitas pelanggan. Berbeda dengan pembelian konvensional, konsumen di *e-commerce* tidak dapat melihat atau mencoba produk secara fisik. Oleh karena itu, calon pembeli sering mengandalkan ulasan pelanggan (*customer review*) sebagai sumber informasi utama [4]. Ulasan tersebut biasanya mencakup opini, pengalaman, dan penilaian konsumen terhadap suatu produk atau layanan. Ulasan positif dapat meningkatkan kepercayaan dan keputusan pembelian, sedangkan ulasan negatif dapat menurunkan minat pembelian dan bahkan memengaruhi reputasi suatu produk atau layanan [5].

Namun demikian, jumlah ulasan pelanggan terus bertambah secara eksponensial menciptakan tantangan tersendiri bagi pelaku bisnis. Volume data yang besar dan tidak terstruktur membuat analisis manual menjadi tidak efisien dan memakan waktu

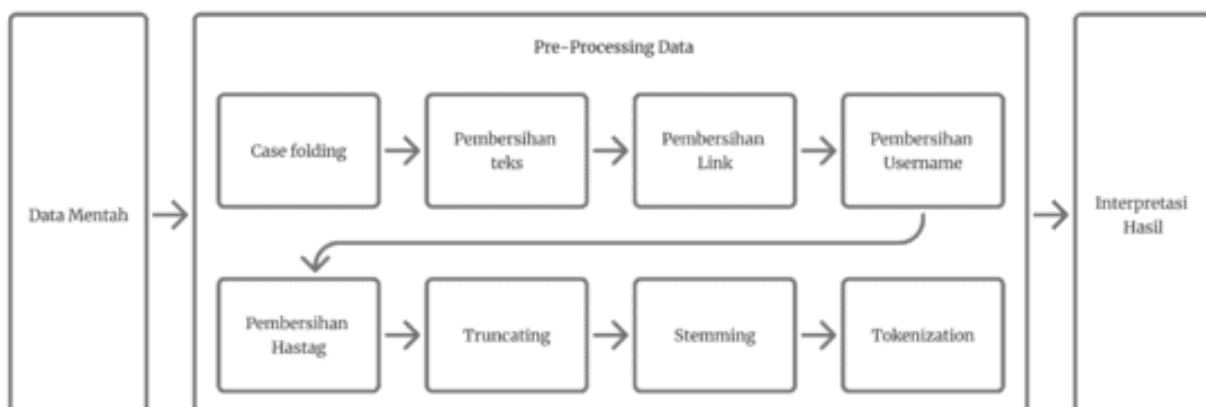
[4]. Selain itu, ulasan pelanggan sering kali mengandung *noise* seperti bahasa informal, singkatan, atau istilah slang yang menyulitkan ekstraksi informasi secara akurat. Tantangan kompleks lainnya mencakup masalah validitas data seperti ulasan palsu (*fake reviews*) atau ulasan yang terlalu subjektif, yang sengaja dibuat untuk memanipulasi penilaian produk [6]. Dalam menghadapi tantangan ini, *text analytics* atau yang dikenal dengan *text mining* dengan pendekatan *Natural Language Processing* (NLP) muncul sebagai solusi yang efektif untuk menganalisis sentimen, pola, dan tren dari ulasan pelanggan secara otomatis. Dengan pendekatan ini, pelaku bisnis dapat mengidentifikasi aspek-aspek produk atau layanan yang paling memengaruhi kepuasan pelanggan [7]. Namun, sebagian besar penelitian terkait analisis sentimen masih berfokus pada bahasa Inggris, sehingga kurang optimal dalam menangkap konteks kalimat dalam bahasa Indonesia. Untuk menjembatani hal ini, model IndoBERT dapat digunakan sebagai solusi karena IndoBERT merupakan model berbasis BERT yang dilatih khusus menggunakan bahasa Indonesia [8].

Relevansi penelitian ini menjadi semakin krusial, mengingat persaingan di pasar *e-commerce* semakin ketat, khususnya pada industri kecantikan [3]. Ulasan pelanggan yang tersebar secara daring tidak hanya membantu calon pembeli dalam mengambil keputusan, tetapi juga membantu perusahaan dalam mengidentifikasi tren pasar [9]. Kemudian, perusahaan juga perlu memahami preferensi dan keluhan pelanggan secara *real-time* melalui ulasan pelanggan agar dapat meningkatkan strategi pemasaran dan pengembangan produk [10]. Selain itu, penerapan *text analytics* juga sejalan dengan tren *data-driven decision making*, di mana keputusan bisnis didasarkan pada analisis data yang akurat dan komprehensif [11].

Berdasarkan tantangan dan peluang yang diuraikan, penelitian ini akan berfokus pada penerapan *text analytics* dengan pendekatan *Natural Language Processing* (NLP) menggunakan IndoBERT untuk menganalisis sentimen ulasan produk *moisturizer* Skintific di Tokopedia. Pemilihan produk ini didasarkan pada empat pertimbangan kunci, yaitu: (1) Skintific merupakan brand lokal unggulan dengan volume transaksi tinggi di kategori *skincare*, (2) *moisturizer* adalah produk yang paling banyak dipakai oleh massa secara umum, (3) produk *moisturizer* termasuk kategori paling banyak diulas di Skintific dan (4) adanya variasi sentimen yang lebar pada ulasan produk ini (antara 1-5 bintang) yang memungkinkan analisis mendalam [12].

2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dan kualitatif dengan metode analisis sentimen berbasis *Natural Language Processing* (NLP) untuk mengklasifikasikan ulasan pelanggan terhadap produk *moisturizer* Skintific di Tokopedia. Model IndoBERT dipilih sebagai dasar pengolahan karena kemampuannya dalam memahami makna dan maksud dari teks berbahasa Indonesia secara mendalam. Kemampuan ini sangatlah relevan, khususnya dalam konteks data yang diperoleh menggunakan Bahasa Indonesia yang beragam dan tidak baku [13]. Data berupa teks ulasan pelanggan yang dikumpulkan melalui teknik *web scraping* dengan bantuan aplikasi ParseHub. ParseHub merupakan alat *web scraping* yang dirancang untuk mengekstrak data dari berbagai jenis situs *web* tanpa adanya pemrograman [14]. Setelah dikumpulkan, data menjalani tahapan *pre-processing* yang meliputi konversi teks ke huruf kecil (*case folding*), pembersihan karakter non-alfabet seperti emoji, penghapusan tautan, *username*, dan *hashtag*, pemotongan teks (*truncate*), *tokenization*, serta reduksi kata melalui *stemming* [15]. Gambar 1 menunjukkan alur dari pemrosesan data pada penelitian ini.

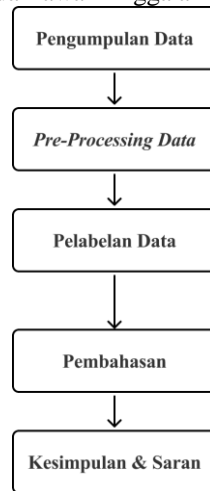


Gambar 1. Alur Pemrosesan Data

Sumber gambar: Data Pribadi

Setelah tahap *pre-processing*, data ulasan yang telah dibersihkan akan digunakan dalam proses pelabelan sentimen menggunakan model IndoBERT. Model ini memproses teks untuk mengidentifikasi opini atau sikap pelanggan terhadap produk, yang dapat dikategorikan menjadi 3 kelas, yaitu positif, negatif, dan netral secara otomatis. Hasil pelabelan ini dianalisis lebih lanjut untuk memahami persepsi konsumen terhadap produk kecantikan yaitu *moisturizer* Skintific di Tokopedia. Tahap akhir penelitian ini adalah penyusunan kesimpulan atas hasil klasifikasi sentimen ulasan, serta saran yang ditujukan untuk penelitian

selanjutnya. Gambar 2 menunjukkan alur penelitian dari awal hingga akhir.



Gambar 2. Alur Penelitian
 Sumber gambar: Dokumen pribadi

3. HASIL PENELITIAN DAN PEMBAHASAN

3.1. Pengumpulan Data

Penelitian ini menggunakan data primer yang diperoleh oleh penulis melalui teknik *web scraping* dengan bantuan aplikasi ParseHub. *Dataset* yang diperoleh terdiri dari 1.394 ulasan produk dari tiga varian produk *moisturizer* yang dijual oleh Skintific di platform Tokopedia, yang mencakup *Ceramide Barrier Moisture Gel*, *Niacinamide Brightening Moisturizer*, dan *Sensitive Moisture Gel*. Proses pengambilan data dilakukan secara acak berdasarkan *rating* (bintang satu sampai dengan bintang lima) untuk memastikan representasi yang seimbang dari berbagai macam sentimen konsumen. Tabel 1 menunjukkan contoh *dataset* yang diperoleh melalui proses *web scraping*.

Tabel 1. Data Ulasan

No.	Ulasan
1	bahannya ringan, gampang diserap kulit, bagus buat melembabkan, sejauh ini beli produk pelembab skintific cocok di kulit aku, gak bikin beruntusan
2	Produk sesuai pesanan, harga juga masih terjangkau, packing aman pake bubble wrap dan juga box, pengiriman lumayan cepat..
3	awalnya maju mundur bingung mau yang biru atau pink, trus akhirnya biru dulu aja. ternyata emang beneran enak banget dipake, ngemoist tapi ga berat di muka. akhirnya ga kerasa kalo pake moisturizer gitu. mana pas co dapet harga murah banget karna pake voucher 🧧🧧🧧
...	...
1394	Kendala: Kemasan tidak aman

Sumber: Tokopedia

3.2. Pre-Processing Data

Tahap *pre-processing data* merupakan proses awal yang bertujuan untuk membersihkan dan menyiapkan data teks agar dapat dianalisis secara optimal oleh model. Dalam penelitian ini, *pre-processing* dilakukan melalui beberapa tahapan, yaitu *case folding*, *data cleaning*, *truncate*, *stemming*, dan *tokenization*. Gambar 3 menunjukkan alur dari tahapan *pre-processing data* pada penelitian ini.



Gambar 3. Alur Pemrosesan Data dari Tahapan Pre-Processing
 Sumber gambar: Dokumen pribadi

Langkah pertama adalah *case folding*, yaitu proses mengubah seluruh huruf dalam teks menjadi huruf kecil (*lowercase*). Selanjutnya, dilakukan *data cleansing* atau pembersihan teks untuk menghapus karakter tidak relevan seperti angka, spasi

berlebih, emoji, link, *username*, simbol seperti *hashtag*, dan mengubah *new line* menjadi spasi. Setelah itu, dilakukan proses *truncate* untuk memotong teks menjadi maksimum 150 karakter pertama yang bertujuan untuk mengurangi panjang data yang berlebihan. Setelah teks dibersihkan dan dipotong, selanjutnya dilakukan proses *stemming*, yaitu mengubah kata-kata dalam teks ke bentuk dasarnya, misalnya kata “diserap” menjadi “serap”, kata “harganya” menjadi “harga”, dan seterusnya. Terakhir, dilakukan *tokenization* untuk memisahkan teks menjadi unit-unit kata yang mana hal ini merupakan dasar dalam representasi vektor teks untuk proses pelatihan model selanjutnya. Tabel 2 menunjukkan hasil dari tahapan *pre-processing data* yang telah diterapkan.

Tabel 2. Pre-Processing Data

No.	Original	Lower	Clean	Truncate	Stemmed	Tokenized
1	bahannya ringan, gampang diserap kulit, bagus buat melembabkan, sejauh ini beli produk pelembab skintific cocok di kulit aku, gak bikin beruntusan	bahannya ringan, gampang diserap kulit, bagus buat melembabkan, sejauh ini beli produk pelembab skintific cocok di kulit aku, gak bikin beruntusan	bahannya ringan gampang diserap kulit bagus buat melembabkan sejauh ini beli produk pelembab skintific cocok di kulit aku gak bikin beruntusan	bahannya ringan gampang diserap kulit bagus buat melembabkan sejauh ini beli produk pelembab skintific cocok di kulit aku gak bikin beruntusan	bahan ringan gampang serap kulit bagus buat melembabkan jauh ini beli produk pelembab skintific cocok di kulit aku gak bikin beruntus	bahan ringan gampang serap kulit bagus buat melembabkan jauh ini beli produk pelembab skintific cocok di kulit aku gak bikin beruntus
2	Produk sesuai pesanan, harga juga masih terjangkau, packing aman pake bubble wrap dan juga box, pengiriman lumayan cepat..	produk sesuai pesanan, harga juga masih terjangkau, packing aman pake bubble wrap dan juga box, pengiriman lumayan cepat..	produk sesuai pesanan harga juga masih terjangkau packing aman pake bubble wrap dan juga box pengiriman lumayan cepat	produk sesuai pesanan harga juga masih terjangkau packing aman pake bubble wrap dan juga box pengiriman lumayan cepat	produk sesuai pesan harga juga masih jangkau packing aman pake bubble wrap dan juga box kirim lumayan cepat	produk sesuai pesan harga juga masih jangkau packing aman pake bubble wrap dan juga box kirim lumayan cepat
3	awalnya maju mundur bingung mau yang biru atau pink, trus akhirnya biru dulu aja. ternyata emang beneran enak banget dipake, ngemoist tapi ga berat di muka. afternya ga kerasa kalo pake moisturizer gitu. mana pas co dapet harga murah banget karna pake voucher👉👉	awalnya maju mundur bingung mau yang biru atau pink, trus akhirnya biru dulu aja. ternyata emang beneran enak banget dipake, ngemoist tapi ga berat di muka. afternya ga kerasa kalo pake moisturizer gitu. mana pas co dapet harga murah banget karna pake voucher👉👉	awalnya maju mundur bingung mau yang biru atau pink trus akhirnya biru dulu aja ternyata emang beneran enak banget dipake ngemoist tapi ga berat di muka afternya ga kerasa kalo pake moisturizer gitu mana pas co dapet harga murah banget karna pake voucher	awalnya maju mundur bingung mau yang biru atau pink trus akhirnya biru dulu aja ternyata emang beneran enak banget dipake ngemoist tapi ga berat di muka	awal maju mundur bingung mau yang biru atau pink trus akhir biru dulu aja nyata emang beneran enak banget dipake ngemoist tapi ga berat di muka	awal maju mundur bingung mau yang biru atau pink trus akhir biru dulu aja nyata emang beneran enak banget dipake ngemoist tapi ga berat di muka
...
1394	Kendala: Kemasan tidak aman	kendala: kemasan tidak aman	kendala kemasan tidak aman	kendala kemasan tidak aman	kendala kemas tidak aman	kendala kemas tidak aman

Sumber: Data Diolah (2025)

3.3. Pelabelan Data

Setelah data dibersihkan, langkah berikutnya merupakan proses pelabelan sentimen yang dikategorikan menjadi tiga

kelompok utama, yaitu positif, netral, dan negatif. Proses pelabelan sentimen digunakan secara otomatis menggunakan model *deep learning* berbasis bahasa Indonesia yaitu model IndoBERT. Model ini mampu mengenali makna dalam ulasan konsumen dan mengklasifikasikannya sesuai dengan nada atau kecenderungan isi ulasan. Setiap ulasan diberi label sentimen berdasarkan analisis isi teksnya. Misalnya, ulasan yang mengandung kata-kata pujian terhadap kualitas produk cenderung diklasifikasikan sebagai kategori positif, sedangkan ulasan yang mengandung keluhan atau kritik dikategorikan sebagai negatif. Adapun ulasan netral biasanya berupa komentar informatif tanpa menyatakan penilaian yang kuat. Proses pelabelan ini dilakukan untuk mempercepat analisis serta mengurangi kemungkinan bias atau kesalahan yang dapat terjadi jika klasifikasi dilakukan secara manual. Tabel 3 merupakan tabel dari hasil ulasan yang sudah diberikan label sentimen.

Tabel 3. Pelabelan Data

No.	Data Ulasan	Label Sentimen
1	bahan ringan gampang serap kulit bagus buat melembabkan jauh ini beli produk pelembab skintific cocok di kulit aku <i>gak</i> bikin beruntus	Positif
2	produk sesuai pesan harga juga masih jangkau packing aman pake bubble wrap dan juga box kirim lumayan cepat	Positif
3	awal maju mundur bingung mau yang biru atau pink trus akhir biru dulu aja nyata emang beneran enak banget dipake ngemoist tapi ga berat di muka afternya ga rasa kalo pake moisturizer gitu mana pas co dapet harga murah banget karna pake voucher	Positif
...
1394	kendala kemasan tidak aman	Negatif

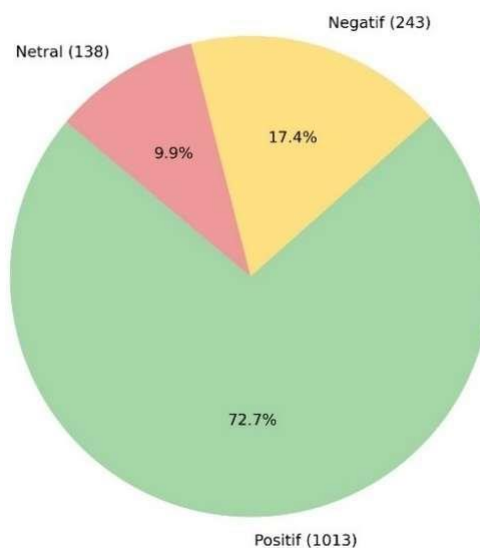
Sumber: Data Diolah (2025)

3.4. Analisis Sentimen Data Olahan

Setelah melalui proses pembersihan data dan pengolahan menggunakan model IndoBERT, hasil klasifikasi sentimen memungkinkan untuk dilakukan analisis secara lebih mendalam. Berdasarkan hasil klasifikasi tersebut, analisis sentimen dalam penelitian ini dapat dibagi ke dalam tiga kategori utama, yaitu: (1) analisis sentimen secara keseluruhan, (2) analisis sentimen berdasarkan jenis produk, dan (3) analisis sentimen berdasarkan *rating* bintang (dari bintang 1 hingga bintang 5).

Berdasarkan hasil klasifikasi sentimen menggunakan model IndoBERT terhadap 1.394 ulasan produk kecantikan yang diambil dari *platform* Tokopedia. Secara keseluruhan, distribusi data ulasan yang telah diberikan label berdasarkan hasil prediksi model IndoBERT menunjukkan bahwa sentimen positif mendominasi dengan persentase sebesar 72,7% atau 1.013 ulasan dari total 1.394 ulasan. Visualisasi distribusi sentimen tersebut ditampilkan dalam bentuk diagram *pie* seperti yang ditunjukkan pada Gambar 3.

Distribusi Sentimen Ulasan Moisturizer Skintific di Tokopedia



Gambar 4. Distribusi Sentimen Ulasan Moisturizer Skintific di Tokopedia

Sumber gambar: Dokumen pribadi

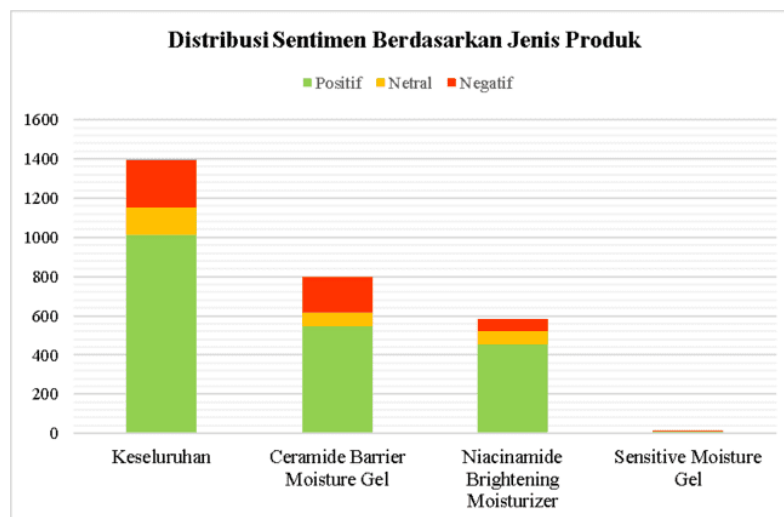
Tingginya proporsi ulasan positif ini mencerminkan bahwa secara keseluruhan, konsumen memiliki opini yang positif

atau baik terhadap produk *moisturizer* Skintific yang dijual melalui *platform* tersebut. Temuan ini mengindikasikan bahwa produk jenis *moisturizer* tersebut secara umum mampu memenuhi ekspektasi dan kebutuhan pengguna, baik dari segi kualitas, kemasan, maupun efektivitasnya selama penggunaan.

Dari distribusi data ulasan, ditemukan ulasan negatif sebesar 17,4% atau 243 ulasan dan ulasan netral sebanyak 9,9%, atau 138 ulasan dari total 1.394 ulasan. Berdasarkan data yang diolah, ulasan yang dikategorikan negatif umumnya berkaitan dengan isu pengemasan, pengiriman, atau ketidaksesuaian produk. Sedangkan ulasan netral umumnya berkaitan dengan komentar informatif tanpa ekspresi emosional yang kuat, seperti pernyataan deskriptif mengenai isi paket atau waktu pengiriman.

Distribusi sentimen ulasan secara keseluruhan memberikan landasan awal untuk memahami persepsi konsumen secara agregat. Berdasarkan distribusi ini, analisis lebih lanjut dapat dilakukan untuk menggali persepsi konsumen secara lebih spesifik, baik dari sisi jenis produk maupun rating bintang yang diberikan. Pendekatan ini memungkinkan identifikasi pola sentimen yang lebih mendalam pada masing-masing kategori, serta memberikan wawasan yang lebih kontekstual terhadap faktor-faktor yang memengaruhi kepuasan atau ketidakpuasan konsumen.

Untuk memperoleh pemahaman yang lebih spesifik mengenai persepsi konsumen, data ulasan kemudian diklasifikasikan berdasarkan jenis produk. Data ulasan dikategorikan berdasarkan tiga jenis produk, yakni *Ceramide Barrier Moisture Gel*, *Niacinamide Brightening Moisturizer*, dan *Sensitive Moisture Gel*. Pembagian ini bertujuan untuk melihat apakah terdapat perbedaan pola sentimen antar varian produk yang ditawarkan. Gambar 5 menunjukkan distribusi sentimen yang beragam pada masing-masing produk.



Gambar 5. Distribusi Sentimen Berdasarkan Jenis Produk
Sumber gambar: Dokumen pribadi

Berdasarkan data ulasan yang dikumpul, produk *Ceramide Barrier Moisture Gel* mendominasi dengan total 799 ulasan. Distribusi sentimennya mencakup 68,46% positif, 9,01% netral, dan 22,53% negatif. Meskipun mayoritas ulasan bersifat positif, produk ini memiliki persentase ulasan negatif tertinggi di antara ketiga produk yang dianalisis. Sementara ini, produk *Niacinamide Brightening Moisturizer* memperoleh 583 ulasan dengan distribusi sentimen yang lebih menguntungkan, yakni 78,22% positif, 11,32% netral, dan hanya 10,46% negatif. Dengan proporsi ulasan positif tertinggi dan negatif terendah, produk ini menonjol sebagai produk yang paling diminati dengan tingkat kepuasan konsumen tertinggi di antara ketiga produk tersebut. Untuk produk *Sensitive Moisture Gel*, jumlah ulasannya jauh lebih sedikit, yakni 12 ulasan dengan 83,33% bersentimen positif dan 16,67% negatif, tanpa adanya ulasan netral. Meskipun menunjukkan indikasi sentimen positif yang cukup kuat, jumlah data yang terbatas belum dapat dianggap sebagai representatif untuk menarik kesimpulan secara menyeluruh.

Selain berdasarkan jenis produk, analisis juga dilakukan dengan mengamati hubungan antara ulasan dan *rating* bintang yang diberikan oleh pengguna. Tujuan dari analisis ini adalah untuk mengevaluasi konsistensi antara penilaian kuantitatif (*rating*) dan penilaian kualitatif (sentimen ulasan), serta meneliti apakah *rating* rendah secara umum mencerminkan sentimen negatif, dan demikian pula sebaliknya. Berdasarkan ulasan dengan *rating* tinggi bintang 5 dan bintang 4, didapatkan bahwa sentimen positif mendominasi secara signifikan dibandingkan kategori sentimen lainnya, yakni sebesar 85,86%, diikuti oleh sentimen netral sebesar 9,59%, dan sentimen negatif sebesar 4,55%. Berbanding terbalik dengan *rating* bintang 5, pada *rating* bintang 1 didapatkan bahwa sentimen negatif mendominasi secara signifikan dibandingkan dengan kategori sentimen lainnya, yakni sebesar 85,96%, diikuti oleh sentimen netral sebesar 8,77% dan sentimen positif sebesar 5,26%. Tabel 4 menunjukkan hasil keseluruhan dari sentimen yang dikategorikan berdasarkan bintang.

Tabel 4. Distribusi Sentimen Berdasarkan Jenis Produk

Hasil Keseluruhan			
Bintang	Positif	Netral	Negatif
5	85,86%	9,59%	4,55%
4	51,17%	12,11%	36,72%
3	13,73%	7,84%	78,43%
2	21,05%	5,26%	73,68%
1	5,26%	8,77%	85,96%

Sumber: Data Diolah (2025)

Secara keseluruhan, hasil analisis ini memperlihatkan adanya keterkaitan yang jelas antara sentimen dalam ulasan dengan *rating* bintang yang diberikan pengguna, dimana *rating* tinggi cenderung diikuti oleh sentimen positif, sementara *rating* rendah secara konsisten menunjukkan dominasi sentimen negatif. Pola ini tidak hanya terlihat pada tingkat agregat, tetapi juga konsisten pada masing-masing jenis produk yang dianalisis.

Temuan ini menegaskan bahwa nilai *rating* dapat dijadikan sebagai indikator awal untuk memprediksi kecenderungan sentimen konsumen, dan sebaliknya, analisis sentimen dapat menjadi alat verifikasi terhadap validitas penilaian pengguna. Dengan demikian, pendekatan ini memberikan nilai tambah dalam upaya perusahaan memahami persepsi pelanggan secara menyeluruh dan *real-time*, serta dapat dijadikan dasar dalam pengembangan strategi pemasaran, peningkatan layanan, dan evaluasi kualitas produk secara lebih akurat.

4. KESIMPULAN

Berdasarkan hasil penelitian mengenai analisis sentimen terhadap ulasan produk kecantikan, yaitu *moisturizer* Skintific di Tokopedia, dengan menggunakan model IndoBERT, dapat disimpulkan bahwa model ini mampu melakukan klasifikasi sentimen dengan baik. IndoBERT menunjukkan performa yang efektif dalam memahami konteks bahasa Indonesia yang digunakan oleh konsumen dalam menyampaikan ulasan mereka. Hasil analisis menunjukkan bahwa sebagian besar ulasan terhadap produk *moisturizer* Skintific didominasi oleh sentimen positif, yakni sebesar 72,7% atau sebanyak 1.013 ulasan dari total 1.394 ulasan yang dianalisis. Dari ketiga varian yang diteliti, yaitu *Ceramide Barrier Moisture Gel*, *Niacinamide Brightening Moisturizer*, dan *Sensitive Moisture Gel*, menunjukkan bahwa varian *Niacinamide Brightening Moisturizer* memiliki proporsi sentimen positif tertinggi. Sementara itu, varian *Ceramide Barrier Moisture Gel* menunjukkan proporsi sentimen negatif yang lebih besar dibandingkan dua varian lainnya, meskipun secara keseluruhan masih didominasi oleh sentimen positif. Selain itu, hasil analisis berdasarkan *rating* bintang menunjukkan adanya konsistensi antara penilaian kuantitatif dan kualitatif. *Rating* bintang lima umumnya disertai dengan ulasan positif, sedangkan *rating* bintang satu cenderung diikuti oleh ulasan negatif. Temuan ini menunjukkan bahwa pendekatan analisis sentimen berbasis *Natural Language Processing*, khususnya dengan pemanfaatan model IndoBERT, dapat memberikan pemahaman yang lebih mendalam terhadap persepsi konsumen terhadap suatu produk. Dengan demikian, model ini berpotensi menjadi alat bantu yang efektif dalam mendukung pengambilan keputusan berbasis data, baik bagi pelaku industri kecantikan maupun platform *e-commerce*.

5. SARAN

Penelitian ini terbatas kepada implementasi dari IndoBERT dalam melakukan analisis sentimen terhadap *online review* produk kecantikan Skintific. Bagi peneliti yang berminat untuk melakukan studi serupa di masa mendatang, disarankan untuk melakukan proses pembersihan data yang lebih komprehensif, seperti penerapan teknik *lemmatization*, guna meningkatkan kualitas data yang diterima oleh model. Selain itu, pelatihan ulang (*fine-tuning*) model juga direkomendasikan agar akurasi hasil analisis sentimen dapat lebih optimal.

6. UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih yang sebesar-besarnya kepada Bapak Frans Mikael Sinaga, selaku dosen pembimbing mata kuliah Kecerdasan Bisnis dan Analitik, atas bimbingan, saran, dan kritik konstruktif yang diberikan selama proses penelitian. Penulis juga ingin menyampaikan apresiasi kepada rekan-rekan di Program Studi Sistem Informasi, Universitas Pelita Harapan yang telah memberikan dukungan, diskusi ilmiah, dan motivasi yang turut berperan dalam kelancaran penyusunan penelitian ini. Selain itu, penulis juga ingin mengucapkan terima kasih kepada keluarga dan teman-teman yang selalu memberikan dukungan moral yang senantiasa menguatkan selama proses penelitian. Segala bentuk dukungan tersebut menjadi kekuatan berarti bagi penulis untuk menyelesaikan penelitian ini.

7. REFERENSI

- [1] M. Praseptiawan, M. O. N. Gultom, and M. C. Untoro, "The Evaluation of E-Commerce Using the Customer Satisfaction Index and Importance Performance Analysis," *Jurnal Sisfokom (Sistem Informasi Dan Komputer)*, vol. 11, no. 1, pp. 60–65, 2022.
- [2] Y. Garamba and A. Nur, "Analisis Kepuasan Pelanggan pada E-Commerce," *Jurnal Ilmu Komputer dan Informatika| E-ISSN: 3063-9026*, vol. 1, no. 2, pp. 27–37, 2024.
- [3] McKinsey & Co., "The beauty market in 2023: A special State of Fashion report," McKinsey & Co. [Online]. Available: <https://www.mckinsey.com/industries/retail/our-insights/the-beauty-market-in-2023-a-special-state-of-fashion-report>
- [4] Y. Liu, Y. Wan, X. Shen, Z. Ye, and J. Wen, "Product customer satisfaction measurement based on multiple online consumer review features," *Information*, vol. 12, no. 6, p. 234, 2021.
- [5] A. N. Pramesti and R. F. Abdillah, "Dampak Rating dan Ulasan Konsumen terhadap Keputusan Pembelian di E-Commerce," in *Prosiding Seminar Nasional Amikom Surakarta*, 2024, pp. 1480–1494.
- [6] W. ShengYi and A. Pandey, "The Impact of Consumer Reviews on Consumer Purchasing Decisions," *St. Theresa Journal of Humanities and Social Sciences*, vol. 9, no. 2, pp. 16–30, 2023.
- [7] E. H. Muktafin, K. Kusri, and E. T. Luthfi, "Analisis Sentimen pada Ulasan Pembelian Produk di Marketplace Shopee Menggunakan Pendekatan Natural Language Processing," *Jurnal Eksplora Informatika*, vol. 10, no. 1, pp. 32–42, 2020.
- [8] Y. A. Singgalen, "IndoBERT-Based Sentiment Analysis for Understanding Hotel Guests' Preferences," *Article in Journal of Computer System and Informatics*, vol. 6, no. 2, pp. 508–520, 2025, doi: 10.47065/josyc.v6i2.6864.
- [9] T. Hou, B. Yannou, Y. Leroy, and E. Poirson, "Mining customer product reviews for product development: A summarization process," *Expert Syst Appl*, vol. 132, pp. 141–150, Oct. 2019, doi: 10.1016/j.eswa.2019.04.069.
- [10] N. Susi, S. Sugiana, and B. Musty, "Analisis Data Sistem Informasi Monitoring Marketing; Tools Pengambilan Keputusan Strategic," *Jutisi : Jurnal Ilmiah Teknik Informatika dan Sistem Informasi*, vol. 12, no. 2, pp. 696–708, 2023.
- [11] J. Noor and W. Shah, "Enhancing Data-Driven Decision Making: AI-Powered Approaches to Text Mining and Knowledge Graphs," *Researchgate*, 2025.
- [12] Tokopedia, "Skintific Official Store," Tokopedia. [Online]. Available: <https://www.tokopedia.com/skintific/>
- [13] Y. A. Singgalen, "Performance Analysis of IndoBERT for Sentiment Classification in Indonesian Hotel Review Data," *Article in Journal of Information System Research*, vol. 6, no. 2, pp. 978–988, Jan. 2025, doi: 10.47065/josh.v6i2.6505.
- [14] M. Zhekova and E. Yumer, "JavaScript Web Scraping Tool for Extraction Information from Agriculture Websites," in *BIO Web of Conferences*, EDP Sciences, 2024.
- [15] F. Koto, A. Rahimi, J. H. Lau, and T. Baldwin, "IndoLEM and IndoBERT: A benchmark dataset and pre-trained language model for Indonesian NLP," *arXiv preprint arXiv:2011.00677*, 2020.