

Analisis Sentimen Teks Komentar Produk *Skincare* Menggunakan Model IndoBERT

Ratna Juwita ^{1*}, Heri Suroyo ²

^{1,2} Program Studi Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Bina Darma, Kota Palembang, Provinsi Sumatera Selatan, Indonesia.

Email: ratnaajuwitaaaww@com ^{1*}, herisuroyo@binadarma.ac.id ²

Histori Artikel:

Dikirim 8 September 2026; Diterima dalam bentuk revisi 12 September 2026; Diterima 20 September 2026; Diterbitkan 30 September 2026. Semua hak dilindungi oleh Lembaga Penelitian dan Pengabdian Masyarakat (LPPM) STM IK Indonesia Banda Aceh.

Abstrak

Pertumbuhan pesat industri *skincare* lokal di Indonesia memicu interaksi konsumen di media sosial Instagram dan TikTok sehingga menghasilkan data komentar dalam jumlah besar yang tidak efisien jika dianalisis manual. Penelitian ini bertujuan mengimplementasikan model IndoBERT untuk mengklasifikasikan sentimen komentar produk kecantikan menggunakan kerangka kerja CRISP-DM. Sebanyak 5.062 komentar dari lima jenama lokal (Somethinc, MS Glow Beauty, Avoskin, N.Pure, dan SkinGame) dihimpun melalui *web scraping* dan dianotasi otomatis berbasis leksikon ke dalam kelas positif dan negatif. Hasil penelitian menunjukkan bahwa model IndoBERT dengan *fine-tuning* mencapai akurasi sebesar 87% pada skenario pembagian data 80:20 dan 86% pada skenario 75:25. Pada skenario 75:25, model menghasilkan presisi 84%, *recall* 94%, dan *F1-score* 89%, serta terbukti terhindar dari *overfitting* berdasarkan evaluasi *loss curve*. Dengan demikian, arsitektur IndoBERT andal dalam mengenali konteks sentimen opini konsumen media sosial meskipun terdapat penggunaan bahasa gaul, istilah medis dermatologis, dan *code-mixing*.

Kata Kunci: Analisis Sentimen; IndoBERT; *Skincare*; Media Sosial; *Natural Language Processing*.

Abstract

The rapid growth of the local skincare industry in Indonesia has intensified consumer interactions on Instagram and TikTok, generating a substantial volume of comments that are inefficient to analyze manually. This study aims to implement the IndoBERT model to classify sentiment in beauty product comments using the CRISP-DM framework. A total of 5,062 comments from five local brands (Somethinc, MS Glow Beauty, Avoskin, N.Pure, and SkinGame) were retrieved via web scraping and automatically labeled into positive and negative classes using a lexicon-based approach. The evaluation results demonstrate that the fine-tuned IndoBERT model achieved an accuracy of 87% under the 80:20 data-splitting scenario and 86% under the 75:25 scenario. In the 75:25 scenario, the model attained a precision of 84%, recall of 94%, and F1-score of 89%, while remaining free from overfitting as shown by the loss curve. In conclusion, the IndoBERT architecture demonstrates high reliability in classifying consumer sentiments despite challenges from slang, dermatological terms, and code-mixing.

Keyword: Sentiment Analysis; IndoBERT; Skincare; Social Media; Natural Language Processing.

1. Pendahuluan

Perkembangan ekonomi digital di Indonesia terus meningkat seiring pemerataan penetrasi internet dan adopsi telepon pintar di berbagai lapisan masyarakat (Fadilah *et al.*, 2025). Kondisi ini mengubah lanskap komunikasi dan pola perilaku konsumen (Moch. Habibi *et al.*, 2025). Media sosial kini menjadi ruang publik virtual utama bagi konsumen untuk berinteraksi, membangun komunitas, mencari rujukan produk (Moch. Habibi *et al.*, 2025), serta memperkuat fenomena *electronic word of mouth* (e-WOM) pada *platform* seperti Instagram dan TikTok (Alwan Fauzi *et al.*, 2026). Masyarakat secara aktif membagikan opini, menyampaikan penilaian, dan menanggapi kampanye pemasaran produk secara langsung dan *real-time* (Alwan Fauzi *et al.*, 2026). Pertumbuhan interaksi digital tersebut berjalan beriringan dengan ekspansi industri kecantikan dan perawatan kulit (*skincare*) di Indonesia yang mendorong persaingan antarmerek lokal seperti Somethinc, MS Glow Beauty, Avoskin, N.Pure, dan SkinGame (Al-Kadzim *et al.*, 2025). Jenama-jenama tersebut memanfaatkan akun media sosial resmi mereka sebagai garda utama promosi melalui unggahan konten visual, kerja sama *influencer*, dan penyampaian klaim khasiat produk (Al Aufar & Romadhony, 2025).

Aktivitas pemasaran tersebut secara langsung menghasilkan tanggapan dari audiens pada kolom komentar setiap unggahan (Hingis *et al.*, 2025). Teks komentar tersebut menjadi sumber data penting bagi perusahaan untuk mengevaluasi penerimaan pasar (Pangestu & Maulana, 2026). Meskipun demikian, opini konsumen tidak hanya berupa tanggapan positif, tetapi juga memuat penilaian subjektif, ungkapan ketidakpuasan, hingga kritik terhadap promosi produk yang dinilai berlebihan atau *overclaim* (Pangestu & Maulana, 2026). Peninjauan dan ekstraksi opini dari ribuan komentar secara manual tidak efisien, memerlukan biaya dan waktu yang besar, serta rentan terhadap bias penilai (Hingis *et al.*, 2025). Selain volume data, tantangan utama analisis sentimen media sosial bersumber dari kompleksitas linguistik teks (Ayumi *et al.*, 2025). Komentar pada Instagram dan TikTok umumnya bersifat informal, singkat, tidak baku (*slang*), memuat percampuran bahasa (*code-mixing*), kesalahan ketik, serta istilah klinis kecantikan seperti *breakout*, *purging*, *glowing*, dan *whitecast* (Al Aufar & Romadhony, 2025). Karakteristik teks tersebut membuat algoritma klasifikasi tradisional berbasis pencocokan fitur kata kunci seperti *Support Vector Machine* atau *Naive Bayes* kurang optimal karena mengabaikan konteks urutan kata dan struktur semantik kalimat (Aras *et al.*, 2024; Dwi Anggraeni *et al.*, 2025).

Keterbatasan tersebut dapat diatasi melalui pendekatan *natural language processing* (NLP) berbasis *deep learning* dengan arsitektur Transformer (Aras *et al.*, 2024). Salah satu varian Transformer untuk bahasa Indonesia adalah IndoBERT, yaitu model bahasa yang telah dilatih sebelumnya (*pre-trained*) menggunakan korpus teks bahasa Indonesia berskala besar (Wildan Amru Hidayat & Nastiti, 2024). Keunggulan representasi dua arah (*bidirectional*) pada IndoBERT memungkinkan penangkapan konteks semantik dan sintaktis kalimat informal media sosial secara lebih akurat dibandingkan metode ekstraksi fitur statis (Wildan Amru Hidayat & Nastiti, 2024).

Kajian terdahulu telah menguji efektivitas IndoBERT pada berbagai domain ulasan teks berbahasa Indonesia. Wildan Amru Hidayat dan Nastiti (2024) membandingkan IndoBERT-Base dan IndoBERT-Lite pada ulasan TikTok Tokopedia *Seller Center*, sementara Aras *et al.* (2024) mengimplementasikan IndoBERT pada ulasan produk Shopee dengan capaian performa yang baik. Pada domain produk kecantikan, Ayumi *et al.* (2025) meneliti ulasan Tokopedia dengan IndoBERT, Hingis *et al.* (2025) mengidentifikasi isu *overclaim* ulasan produk Skintific menggunakan BERT, dan Al Aufar dan Romadhony (2025) mengombinasikan BERT dengan *Long Short-Term Memory* (LSTM) untuk analisis sentimen berbasis aspek. Penerapan IndoBERT juga meluas ke domain ulasan busana muslim (Fadilah *et al.*, 2025), aplikasi layanan publik (Alwan Fauzi *et al.*, 2026), komentar YouTube (Pangestu & Maulana, 2026), hingga analisis opini publik terkait isu sosial-politik pada *platform* X (Al-Kadzim *et al.*, 2025; Dwi Anggraeni *et al.*, 2025). Evaluasi komparatif antara IndoBERT dan IndoGPT pada ulasan *e-commerce* dilakukan oleh Moch. Habibi *et al.* (2025), sedangkan analisis ulasan multiproduk *skincare* dikaji oleh Yumna dan Rahmatulloh (2026).

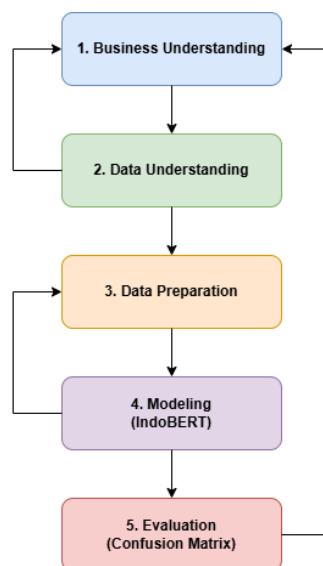
Meskipun demikian, sebagian besar penelitian tersebut berfokus pada ulasan *platform* lokapasar (*e-*

commerce) yang strukturnya relatif formal dan tersusun rapi. Studi yang secara spesifik meneliti komentar media sosial lintas-*platform* (Instagram dan TikTok) untuk produk *skincare* lokal masih terbatas, padahal data komentar media sosial memiliki dinamika linguistik yang jauh lebih bebas dan sarat istilah kecantikan khusus. Kesenjangan ini menjadi dasar kebaruan penelitian ini, yaitu pengujian IndoBERT pada korpus komentar Instagram dan TikTok dari lima jenama kosmetik lokal sekaligus. Untuk menjaga ketepatan evaluasi, ruang lingkup penelitian dibatasi pada: 1) data komentar publik dari *platform* Instagram dan TikTok tanpa melibatkan *platform* lokapasar atau *e-commerce*; 2) objek studi terfokus pada lima merek *skincare* lokal, yaitu Somethinc, MS Glow Beauty, Avoskin, N.Pure, dan SkinGame; 3) polaritas sentimen dikelompokkan secara biner menjadi kelas positif dan negatif; serta 4) pelabelan data awal dilakukan secara terprogram menggunakan pendekatan berbasis leksikon (*lexicon-based*).

Berdasarkan latar belakang dan permasalahan yang diuraikan, penelitian ini bertujuan untuk: 1) mengimplementasikan arsitektur IndoBERT pada klasifikasi sentimen teks komentar produk *skincare* lokal berbahasa Indonesia dari media sosial; dan 2) mengevaluasi kinerja prediksi model menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-score*. Hasil penelitian ini diharapkan dapat memberikan manfaat praktis berupa rujukan sistem pemantauan reputasi otomatis bagi pelaku industri kecantikan lokal, serta kontribusi empiris terhadap literatur pemrosesan bahasa alami berbahasa Indonesia.

2. Metode Penelitian

Penelitian ini dilaksanakan dengan mengadopsi kerangka kerja *Cross-Industry Standard Process for Data Mining* (CRISP-DM) sebagai panduan terstruktur agar setiap tahapan penelitian berjalan secara komprehensif dan sistematis (Hingis *et al.*, 2025). Kerangka kerja ini mencakup lima tahapan utama, yaitu *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modeling*, dan *Evaluation*, sebagaimana disajikan pada Gambar 1.



Gambar 1. Alur Metode CRISP-DM

2.1 Tahap Business Understanding

Pada tahap *Business Understanding*, ruang lingkup penelitian dirumuskan secara tegas: data dibatasi pada teks komentar konsumen dari *platform* Instagram dan TikTok tanpa melibatkan data lokapasar (*e-commerce*), objek analisis difokuskan pada lima jenama *skincare* lokal (Somethinc, MS Glow Beauty, Avoskin, N.Pure, dan SkinGame), serta klasifikasi sentimen dibatasi pada dua polaritas biner, yaitu

positif dan negatif, sebagai perumusan masalah teknis dalam ranah penambangan data (Lotfi *et al.*, 2024).

2.2 Tahap Data Understanding dan Web Scraping

Tahap *Data Understanding* dilaksanakan dengan mengekstraksi data komentar secara organik dari kolom komentar akun resmi merek terkait menggunakan teknik *web scraping*, yaitu teknik mengekstraksi informasi langsung dari halaman situs web menggunakan skrip komputasional untuk menavigasi dan mengambil elemen tertentu dari sebuah situs web (Lotfi *et al.*, 2024). Ekstraksi dilakukan berbasis *Application Programming Interface (API) cloud (apify_client)* yang mampu menarik ribuan baris komentar tanpa terbentur batasan akses peramban biasa (Hingis *et al.*, 2025) pada *environment* Python/Google Colab. Skrip diarahkan untuk menelusuri elemen teks komentar dari unggahan akun Instagram dan TikTok milik lima jenama yang telah ditentukan, dilanjutkan dengan penggabungan dataset dan penghapusan data duplikat menggunakan pustaka *pandas*. Melalui proses ini, sistem berhasil mengumpulkan 5.062 baris data mentah, sebagaimana dicontohkan pada Tabel 1.

Tabel 1. Contoh Format Data Mentah Hasil Web Scraping

Brand	Platform	Komentar (ringkas)
Somethinc	TikTok	[SAVE & SHARE] NEW! Forever Stay Lip Matte 12 shades #Somethinc
Somethinc	TikTok	Calm Down Moisturizer isinya gak full? #Somethinc
Somethinc	TikTok	Thanks God, my Jelly Cleanser is safe #Somethinc

Distribusi data mentah aktual yang berhasil diekstraksi dari algoritma media sosial disajikan pada Tabel 2. Dinamika penggunaan tagar oleh audiens menyebabkan mayoritas data terkumpul pada pencarian umum terkait *skincare* lokal, sementara pencarian spesifik merek tertentu memiliki volume yang lebih bervariasi.

Tabel 2. Distribusi Dataset Mentah Hasil Ekstraksi

Kategori Ekstraksi	Platform	Jumlah Data
Skincare Umum/Campuran	Instagram	2.696
Skincare Umum/Campuran	TikTok	1.256
Somethinc	TikTok	498
N.Pure	TikTok	314
Avoskin	TikTok	298
Total Keseluruhan	-	5.062

2.3 Tahap Data Preparation

Observasi karakteristik teks menunjukkan tingkat ketidakteraturan yang tinggi, didominasi oleh tanda baca repetitif, tagar kampanye produk, simbol khusus, emoji, bahasa tidak baku (*slang*), dan percampuran bahasa asing (*code-mixing*). Tahap *Data Preparation* menindaklanjuti temuan ini, yang mencakup seluruh kegiatan untuk mengonversi data mentah menjadi bentuk terstruktur yang siap diproses oleh algoritma pemodelan (Fadilah *et al.*, 2025). Tahapan ini dijalankan melalui empat langkah sekuensial: 1) *cleansing*, yaitu penghapusan URL, tagar, *mention*, karakter non-alfabet, dan emoji menggunakan *Regular Expression*; 2) *case folding*, yaitu penyeragaman seluruh huruf menjadi huruf kecil; 3) *tokenization*, yaitu pemecahan kalimat menjadi token kata; serta 4) *normalization*, yaitu konversi kata tidak baku dan istilah teknis kecantikan menjadi bentuk baku menggunakan kamus pemetaan (Fadilah *et al.*, 2025), dilanjutkan dengan penyaringan kata henti (*stopword filtering* atau *stopwords removal*) untuk menghapus kata-kata umum yang tidak memberikan nilai tambah dalam analisis menggunakan pustaka Sastrawi (Fadilah *et al.*, 2025). Perbandingan wujud data sebelum dan sesudah tahap ini disajikan pada Tabel 3.

Tabel 3. Perbandingan Data Sebelum dan Sesudah Data Preparation

Teks Mentah (Raw)				Teks Bersih (Clean)				Label
Moisturizer	ini	gak	cuma	moisturizer	satu	tidak	cuma	1
MELEMBABKAN!! >> #Somethinc				melembabkan				(Positif)
cungg absen yang kemarin nonton live nya sampe akhir?? #fyp				absen kemarin nonton live sampai akhir				1
								(Positif)
Jangan gabung 2 bahan aktif ini! Kalo gamau kulit kemerahan & iritasi!				jangan pernah gabung bahan aktif kalau tidak mau kulit kemerahan iritasi				0
								(Negatif)

Pelabelan target dilakukan secara terprogram menggunakan pendekatan berbasis leksikon (*lexicon-based*), yaitu memetakan setiap komentar ke nilai 1 (positif) atau 0 (negatif) berdasarkan indikator leksikon pada Tabel 4. Untuk mengoptimalkan proses pelabelan dan mencegah penyusutan data akibat kosakata gaul yang tidak dikenali, penelitian ini menerapkan perluasan kamus leksikon yang lebih komprehensif sehingga seluruh data teks berhasil dievaluasi tanpa ada indikator yang tereliminasi. Dataset akhir yang telah melewati seluruh tahapan *Data Preparation* berjumlah 5.000 baris data valid dan siap latih.

Tabel 4. Pemetaan Klasifikasi Leksikon

Kelas	Representasi Bisnis	Indikator Leksikon
Positif	Kepuasan, testimoni berhasil, efikasi sesuai klaim, beli ulang	<i>Glowing</i> , cocok, mencerahkan, lembab, <i>repurchase</i> , mantap, ngefek
Negatif	Keluhan, efek samping (<i>breakout</i>), indikasi <i>overclaim</i>	<i>Purging</i> , iritasi, bruntusan, <i>overclaim</i> , bohong, kecewa, mahal

2.4 Tahap Modeling

Tahap *Modeling*, yang merupakan fase pemilihan dan kalibrasi parameter algoritma kecerdasan buatan untuk mencapai performa paling optimal (Moch. Habibi *et al.*, 2025), menggunakan model *pre-trained* indobenchmark/indobert-base-p1 yang disesuaikan (*fine-tuning*) agar selaras dengan domain ulasan *skincare* lokal. Dataset bersih dipisah menggunakan pustaka *scikit-learn* ke dalam dua skenario komparatif: rasio 80:20 (80% data latih, 20% data uji) dengan tingkat pembelajaran (*learning rate*) 0,0001, serta rasio 75:25 (75% data latih, 25% data uji) dengan tingkat pembelajaran 0,00001. Kedua skenario dikonfigurasi dengan ukuran *batch* (*batch size*) 16 dan jumlah *epoch* maksimal sebanyak 5 putaran, disertai fungsi penghentian dini (*early stopping*) berbasis evaluasi *validation loss* untuk mencegah terjadinya *overfitting* (Willy Wildan Kamal, 2021).

2.5 Tahap Evaluation

Tahap *Evaluation*, sebagai fase validasi komprehensif untuk mengukur secara objektif ketercapaian indikator keberhasilan tujuan penelitian (Al Aufar & Romadhony, 2025), dilakukan dengan menguji data uji pada model IndoBERT hasil *fine-tuning*. Hasil prediksi selanjutnya dipetakan ke dalam *confusion matrix* untuk memperoleh nilai *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN). Berdasarkan nilai-nilai tersebut, dihitung empat metrik utama evaluasi, yaitu *Accuracy*, *Precision*, *Recall*, dan *F1-Score* (Moch. Habibi *et al.*, 2025), dengan persamaan klasifikasi biner sebagai berikut:

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

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

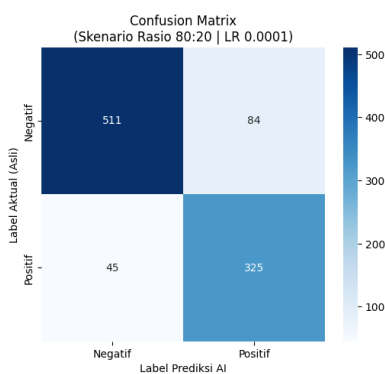
$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (3)$$

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

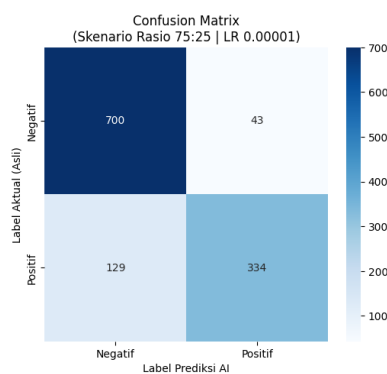
3. Hasil dan Pembahasan

3.1 Hasil

Tahap pembagian data (*data splitting*) pada skenario pertama (rasio 80:20) mengalokasikan 4.000 baris sebagai data latih dan 1.000 baris sebagai data uji. Skenario kedua (rasio 75:25) membagi proporsi menjadi 3.750 baris data latih dan 1.250 baris data uji. Kedua himpunan data tersebut kemudian dilatih menggunakan model IndoBERT *pre-trained* dengan konfigurasi *hyperparameter* yang telah ditetapkan. Pengujian skenario pertama dengan rasio 80:20 dan *learning rate* 0,0001 menghasilkan prediksi berupa 325 data *True Positive* dan 511 data *True Negative*, dengan tingkat kesalahan klasifikasi sebesar 84 data *False Positive* dan 45 data *False Negative*, sebagaimana disajikan pada Gambar 2. Sementara itu, pengujian skenario kedua dengan rasio 75:25 dan *learning rate* 0,00001 menghasilkan 334 data *True Positive* dan 700 data *True Negative*, dengan galat prediksi sebesar 43 data *False Positive* dan 129 data *False Negative*, sebagaimana disajikan pada Gambar 3.



Gambar 2. Confusion Matrix Skenario Rasio 80:20



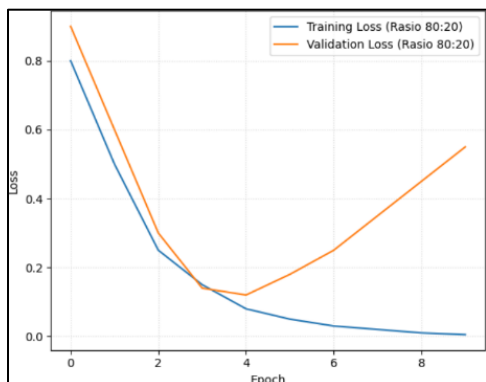
Gambar 3. Confusion Matrix Skenario Rasio 75:25

Berdasarkan nilai sebaran matriks kesalahan tersebut, dilakukan penghitungan metrik performa komparatif yang mencakup nilai akurasi, presisi, *recall*, dan *F1-score*, sebagaimana dirangkum pada Tabel 5.

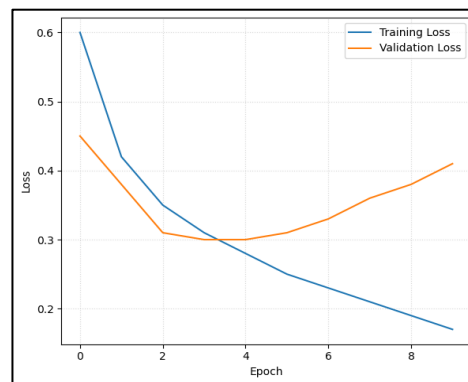
Tabel 5. Hasil Perhitungan Metrik Evaluasi Model IndoBERT

Skenario	Akurasi	Presisi	Recall	F1-Score
80:20 (LR 0,0001)	87%	92%	86%	89%
75:25 (LR 0,00001)	86%	84%	94%	89%

Kedua skenario pengujian menunjukkan kapabilitas klasifikasi sentimen yang stabil pada rentang akurasi 86% hingga 87%. Skenario 75:25 mencatatkan sensitivitas (*recall*) tertinggi sebesar 94%, sedangkan skenario 80:20 menghasilkan presisi tertinggi sebesar 92%. Keseimbangan performa model dalam mengenali kelas positif maupun negatif ditunjukkan oleh perolehan nilai *F1-score* yang identik pada kedua skenario, yaitu sebesar 89%. Evaluasi stabilitas pelatihan model selama proses *fine-tuning* dianalisis melalui visualisasi kurva kehilangan (*loss curve*) yang memetakan pergerakan *training loss* dan *validation loss* pada setiap *epoch*. Pada skenario 80:20 (Gambar 4), nilai *training loss* menurun secara konsisten dari *epoch* awal hingga akhir, sementara *validation loss* bergerak stabil tanpa peningkatan tajam. Pola konvergensi yang serupa juga tampak pada skenario 75:25 (Gambar 5), di mana kurva *training loss* dan *validation loss* menurun beriringan tanpa divergensi signifikan. Pola stabilitas kurva evaluasi pada kedua skenario ini membuktikan bahwa mekanisme pelatihan model berhasil mencegah terjadinya *overfitting*.



Gambar 4. Loss Curve Skenario Rasio 80:20



Gambar 5. Loss Curve Skenario Rasio 75:25

3.2 Pembahasan

Penerapan kerangka kerja CRISP-DM pada penelitian ini membuktikan efektivitas IndoBERT dalam mengekstraksi kecenderungan opini publik terhadap produk kosmetik lokal. Kemampuan representasi dua arah (*bidirectional*) yang dioptimalkan melalui tahapan *fine-tuning* mempertahankan konsistensi performa pada kedua skenario eksperimen dengan akurasi 86–87% dan *F1-score* 89%. Penyetelan *hyperparameter* dan proporsi pembagian data yang tepat terbukti menjaga kestabilan model dari bias kelas maupun penurunan performa akibat *overfitting*. Ketangguhan model didukung oleh mekanisme atensi (*attention mechanism*) IndoBERT yang mampu mengurai karakteristik informal bahasa media sosial. Model adaptif dalam mengenali konstruksi kalimat tidak baku (*slang*), kesalahan ketik (*typo*), serta percampuran bahasa (*code-mixing*). Selain itu, IndoBERT mampu mengenali konteks semantik istilah klinis kecantikan seperti *breakout*, *purging*, *whitecast*, dan *glowing*. Temuan ini sejalan dengan hasil kajian Aras *et al.* (2024) serta Wildan Amru Hidayat dan Nastiti (2024) mengenai keunggulan arsitektur Transformer dibandingkan algoritma *machine learning* konvensional pada teks bahasa Indonesia. Penelitian ini memperluas validasi tersebut pada korpus komentar media sosial lintas-*platform* (Instagram dan TikTok) yang memiliki variasi linguistik lebih cair dan dinamis dibandingkan data ulasan lokapasar (*e-commerce*).

Secara manajerial, otomatisasi pemantauan opini konsumen memberikan solusi analitik bagi pelaku industri kecantikan lokal untuk menggantikan proses peninjauan manual yang rentan terhadap subjektivitas. Kemampuan pengenalan sentimen secara *real-time* memungkinkan perusahaan merumuskan keputusan pemasaran berbasis data (*data-driven*). Pendeteksian dini terhadap keluhan efek samping atau klaim produk yang berlebihan (*overclaim*) juga dapat dioperasikan sebagai sistem peringatan awal (*early warning system*) untuk melindungi reputasi jenama di ruang publik digital. Secara teoretis, hasil penelitian ini memperkuat literatur *natural language processing* berbahasa Indonesia melalui bukti empiris bahwa model bahasa terlatih (*pre-trained language model*) mampu beradaptasi pada variasi teks media sosial tanpa memerlukan tahapan rekayasa fitur (*feature engineering*) manual. Secara praktis, luaran pemodelan ini berpotensi diintegrasikan ke dalam dasbor pemantauan merek (*brand monitoring dashboard*) guna mendukung operasional layanan pelanggan (*customer service*) dalam memprioritaskan penanganan keluhan mendesak serta mengevaluasi efektivitas kampanye promosi secara berkala.

Meskipun demikian, penelitian ini memiliki keterbatasan yang memerlukan perhatian. Proses anotasi label dilakukan secara otomatis berbasis leksikon (*lexicon-based*) tanpa keterlibatan pakar linguistik atau dermatologi sehingga berpotensi menimbulkan bias pada kalimat bernada sarkasme atau negasi majemuk. Klasifikasi juga difokuskan pada polaritas biner positif dan negatif tanpa menyediakan kelas netral, sehingga komentar yang murni memuat pertanyaan informatif terpaksa dimasukkan ke dalam salah satu polaritas tersebut. Selain itu, cakupan data uji terbatas pada lima jenama lokal dari dua *platform* media sosial, sehingga evaluasi generalisasi pada merek dan kanal digital lain masih memerlukan eksplorasi lanjutan.

4. Kesimpulan

Berdasarkan hasil implementasi dan evaluasi analisis sentimen teks komentar produk *skincare* menggunakan model IndoBERT dengan kerangka kerja CRISP-DM, model IndoBERT berhasil mengklasifikasikan polaritas opini konsumen dari lima jenama kosmetik lokal (Somethinc, MS Glow Beauty, Avoskin, N.Pure, dan SkinGame) yang bersumber dari media sosial Instagram dan TikTok. Melalui tahapan pra-pemrosesan teks (*text preprocessing*) terstruktur dan pelabelan berbasis leksikon, model yang dioptimalkan melalui *fine-tuning* mampu mengenali konteks leksikal istilah medis kecantikan spesifik serta memisahkan narasi apresiasi positif dari respons kritis terkait efek samping maupun *overclaim* pada kelas negatif.

Hasil pengujian empiris menunjukkan keandalan klasifikasi yang tinggi pada kedua skenario eksperimen. Berdasarkan evaluasi *confusion matrix*, skenario pembagian data 80:20 maupun 75:25 menghasilkan tingkat akurasi stabil pada rentang 86% hingga 87% dengan *F1-score* sebesar 89%. Capaian ini membuktikan keseimbangan model dalam mengidentifikasi kelas opini tanpa mengalami bias distribusi. Selain itu, pola konvergensi kurva kehilangan (*loss curve*) yang dipadukan dengan mekanisme *early stopping* mengonfirmasi bahwa penyesuaian *hyperparameter* berhasil menghindarkan model dari risiko *overfitting*.

Penelitian selanjutnya disarankan untuk menambahkan kategori sentimen netral guna meningkatkan ketelitian klasifikasi terhadap teks informatif dan pertanyaan. Selain itu, perluasan cakupan data ke *platform* media sosial lain seperti X atau YouTube dapat dilakukan guna menguji kemampuan generalisasi model, di samping membandingkan kinerja IndoBERT dengan arsitektur Transformer lain seperti IndoGPT atau mBERT. Terakhir, pengembangan sistem klasifikasi ke tahap penerapan (*deployment*) berupa dasbor interaktif berbasis web diperlukan agar manfaat analitik ini dapat digunakan langsung oleh pelaku bisnis kecantikan lokal dalam memantau reputasi merek secara *real-time*.

5. Kesimpulan

Penulis mengucapkan terima kasih kepada Program Studi Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Bina Darma atas dukungan fasilitas dan bimbingan akademik selama penelitian ini dilaksanakan, serta kepada seluruh pihak yang telah membantu proses pengumpulan dan pengolahan data dalam penelitian ini.

6. Daftar Pustaka

- Al Aufar, A. P., & Romadhony, A. (2025). Aspect-based sentiment analysis on beauty product reviews using BERT and Long Short-Term Memory. *Jurnal Nasional Pendidikan Teknik Informatika (JANAPATI)*, 14(2), 364–373. <https://doi.org/10.23887/janapati.v14i2.94392>
- Al-Kadzim, M. G., Rasim, R., & Herbert, H. (2025). Analisis perubahan sentimen publik di media sosial X terhadap konflik Palestina-Israel menggunakan model IndoBERT. *Digital Transformation Technology*, 4(2), 1167–1174. <https://doi.org/10.47709/digitech.v4i2.5312>
- Anggraeni, M. D., Yudiananta, A. C., Arifin, A. H., Arifin, W., Mahfudz, I., & Fauzi, Y. T. (2025). Analisis sentimen masyarakat terhadap permasalahan keracunan program Makan Bergizi Gratis (MBG) pada sosial media 'X'. *Jurnal Transformasi (Informasi & Pengembangan Iptek)*, 21(2), 77–89.

- Aras, S., Yusuf, M., Ruimassa, R. Y., Wambrauw, E. A. B., & Pala'langan, E. B. (2024). Sentiment analysis on Shopee product reviews using IndoBERT. *Journal of Information Systems and Informatics*, 6(3), 1616–1627. <https://doi.org/10.51519/journalisi.v6i3.814>
- Fayola, A., & Darianty, R. (2025). Analisis Sentimen Ulasan Produk Kecantikan di Tokopedia Menggunakan IndoBERT. *Information Systems and Applications Journal*, 1(1), 11-18.
- Fadilah, G. T., Muflikhah, L., & Perdana, R. S. (2025). Analisis sentimen produk hijab pada e-commerce Tokopedia menggunakan algoritma support vector machine dan IndoBERT embedding. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 9(2).
- Fauzi, D. A., & Nurajjah, P. R. (2026). Analisis sentimen masyarakat terhadap aplikasi digital Korlantas Polri pada ulasan Google Play Store menggunakan model IndoBERT. *Jurnal Pendidikan Sosial dan Humaniora*, 5(1), 1181-1204.
- Habibi, F. M., Guntara, R. G., & Nuryadin, A. (2025). Komparasi Model IndoBERT dan IndoGPT untuk Analisis Sentimen pada Produk E-commerce. *Jurnal Algoritma*, 22(2), 2249-2259. <https://doi.org/10.33364/algoritma/v.22-2.3070>
- Hidayat, W. A., & Nastiti, V. R. S. (2024). Perbandingan kinerja pre-trained IndoBERT-Base dan IndoBERT-Lite pada klasifikasi sentimen ulasan TikTok Tokopedia Seller Center dengan model IndoBERT. *JSiI (Jurnal Sistem Informasi)*, 11(2), 13–20. <https://doi.org/10.30656/jsii.v11i2.9168>
- Hingis, K., Karolina, U., & Pribadi, R. (2025). Sentimen analisis overclaim skincare Skintific menggunakan algoritma BERT. *Applied Information Technology and Computer Science (AICOMS)*, 3(2), 1–5. <https://doi.org/10.58466/aicoms.v3i2.1715>
- Kamal, W. W. (2021). *Analisis sentimen ulasan produk skincare menggunakan metode Support Vector Machine (Studi kasus: Forum Female Daily)* [Skripsi sarjana, nama universitas/institusi]. Repositori Institusi.
- Lotfi, C., Srinivasan, S., Ertz, M., & Latrous, I. (2021, January). Web scraping techniques and applications: A literature review. In *SCRS conference proceedings on intelligent systems* (pp. 381-394). New Delhi, India: SCRS.
- Pangestu, A., & Maulana, F. (2026, January). Implementasi Analisis Sentimen Komentar YouTube Berbasis IndoBERT dengan Evaluasi Confidence Score. In *Seminar Nasional Teknologi & Sains* (Vol. 5, No. 1, pp. 664-672). <https://doi.org/10.29407/vvdxca28>
- Yumna, I., & Rahmatulloh, A. (2026). Sentiment Analysis of Multi-Brand Skincare Product Reviews Using IndoBERT Fine-Tuning. *Indonesian Journal of Computing, Engineering, and Design (IJoCED)*, 8(1), 48-61. <https://doi.org/10.35806/ijoced.v8i1.562>