Jurnal ICT : Information Communication & Technology
Vol. 26, N0.1, Juli 2026, pp. 175-183
p-ISSN: 2302-0261, e-ISSN: 2303-3363, DOI: https://doi.org/10.36054/jict-ikmi.v26i1.369

Implementasi XLM-RoBERTa untuk Multilingual Aspect-Based Sentiment Analysis (Studi Kasus Indomie Korean Ramyeon Series)

Moch. Yusuf Haidar Ali Ramdhani[1*], Rangga Gelar Guntara[2], Muhammad Rizki Nugraha[3]
[1,2,3]Program Studi Bisnis Digital, Universitas Pendidikan Indonesia, Indonesia
Email: [1]mochyar@upi.edu, [2] ranggagelar@upi.edu, [3]murinu@upi.edu

INFORMASI ARTIKEL

Histori artikel:

Naskah masuk, 10 Juni 2026
Direvisi, 25 Juni 2026
Diiterima, 24 Juli 2026

Kata Kunci:

XLM-RoBERTa
Aspect-Based Sentiment Analysis
Indomie Korean Ramyeon Series

ABSTRAK

Abstract- *The growth of social media has led to an increase in the volume of public responses, making Aspect-Based Sentiment Analysis (ABSA) a relevant approach for understanding public opinion at a granular level. However, monolingual models often struggle with multilingual social media data that is rich in extreme code-mixing phenomena. This study implements a Twitter-pre-trained XLM-RoBERTa model. Using the CRISP-DM framework, a dataset was collected in Indonesian, English, and Korean regarding public opinion on the Indomie Korean Ramyeon Series products on the X platform. This study employs a hybrid approach, in which aspect category extraction is performed using a rule-based method based on a lexical dictionary, while sentiment polarity classification is optimized using a deep learning model. The evaluation results show that this specific preprocessing workflow is far superior in terms of overall accuracy of 83.92%, a Weighted Average F1-score of 0.8385, and a Macro Average F1-score of 0.8241. The impact of class imbalance was successfully mitigated in the Negative class, achieving an F1-score of 0.7934 thanks to adjustments to the prediction error penalty weights. During the deployment phase, predictions were dominated by positive sentiment across all three aspects, with the "Collaboration" aspect recording the highest volume. This approach proved effective in objectively mapping market response dynamics to support the analysis of public preferences and the evaluation of brand collaboration strategies on social media.*

Abstrak- Perkembangan media sosial mendorong peningkatan volume respons publik, sehingga *Aspect-Based Sentiment Analysis* (ABSA) menjadi pendekatan relevan untuk memahami opini secara granular. Namun, model *monolingual* sering gagal menghadapi data media sosial lintas bahasa yang kaya akan fenomena *code-mixing* ekstrem. Penelitian ini mengimplementasikan model XLM-RoBERTa berbasis *pre-training* Twitter. Menggunakan kerangka CRISP-DM, dataset dikumpulkan dalam bahasa Indonesia, Inggris, dan Korea terkait opini publik pada produk Indomie Korean Ramyeon Series di platform X. Penelitian ini menerapkan *hybrid approach*, di mana ekstraksi kategori aspek dieksekusi menggunakan metode *rule-based* berbasis kamus leksikon, sementara klasifikasi polaritas sentimen dioptimalkan menggunakan model *deep learning*. Hasil evaluasi membuktikan alur prapemrosesan khusus ini jauh lebih unggul dengan capaian akurasi keseluruhan sebesar 83,92%, *Weighted Average* F1-score 0,8385, dan *Macro Average* F1-score 0,8241. Dampak *class imbalance* berhasil dimitigasi optimal pada kelas *Negative* dengan capaian F1-score 0,7934 berkat penyesuaian bobot penalti kesalahan prediksi. Pada tahap *deployment*, prediksi didominasi oleh sentimen *Positive* di ketiga aspek, dengan aspek Kolaborasi mencatat volume terbesar. Pendekatan ini terbukti objektif dalam memetakan dinamika respons pasar guna mendukung analisis preferensi publik dan evaluasi strategi kolaborasi merek di media sosial.

Penulis Korespondensi*:*

Rangga Gelar Guntara
Program Studi Bisnis Digital,
Universitas Pendidikan Indonesia
Jalan Dadaha No. 18, Kahuripan, Tawang, Kota Tasikmalaya,
Email: ranggagelar@upi.edu

1. Pendahuluan

Di era ekonomi digital yang berkembang pesat, media sosial telah bertransformasi dari sekadar saluran komunikasi menjadi ruang utama bagi publik untuk menyampaikan opini dan pengalaman konsumsi secara terbuka. Dengan lebih dari 5,4 miliar pengguna media sosial aktif di seluruh dunia pada 2025 [1] , volume percakapan yang dihasilkan setiap detiknya menyimpan nilai strategis yang besar bagi dunia bisnis. Namun, sekitar 90 persen dari data tersebut berbentuk teks tidak terstruktur [2] yang tidak dapat langsung diinterpretasikan oleh sistem analitik konvensional. Sementara riset pasar berbasis survei dinilai terlalu lambat untuk menangkap dinamika publik yang bergerak cepat. Kondisi tersebut mendorong kebutuhan mendesak bagi perusahaan untuk mengadopsi pendekatan berbasis kecerdasan buatan guna mempertahankan keberlanjutan bisnis[3].

Salah satu pendekatan yang relevan adalah analisis sentimen berbasis *Natural Language Processing* (NLP), yang terbukti lebih andal dalam menangkap nuansa linguistik teks dibandingkan metode berbasis leksikon tradisional [4] . Lebih spesifik, *Aspect-Based Sentiment Analysis* (ABSA) memungkinkan identifikasi sentimen pada level aspek tertentu sehingga menghasilkan wawasan yang jauh lebih granular [5].

Fenomena yang menarik untuk diteliti dalam penelitian ini adalah peluncuran Indomie Korean Ramyeon Series oleh PT Indofood CBP Sukses Makmur Tbk, yang diperkuat dengan penunjukan grup K-pop NewJeans sebagai *global brand ambassador* [6] . Langkah strategis ini memicu gelombang percakapan di platform X yang sangat kaya untuk dianalisis. Namun, opini publik terhadap produk ini tidak ditulis dalam satu bahasa yang seragam. Opini terdiri dalam berbagai bahasa, seperti bahasa Indonesia, bahasa Inggris, dan bahasa Korea. Dalam satu teks juga disertai slang K-pop, *noise* khas media sosial, serta *code-mixing*. Berdasarkan tantangan data tersebut, model NLP *monolingual* seperti IndoBERT terbukti kurang optimal menghadapi fenomena *code-switching* [7], sementara pendekatan translasi otomatis justru merusak konteks dan nuansa asli teks [8].

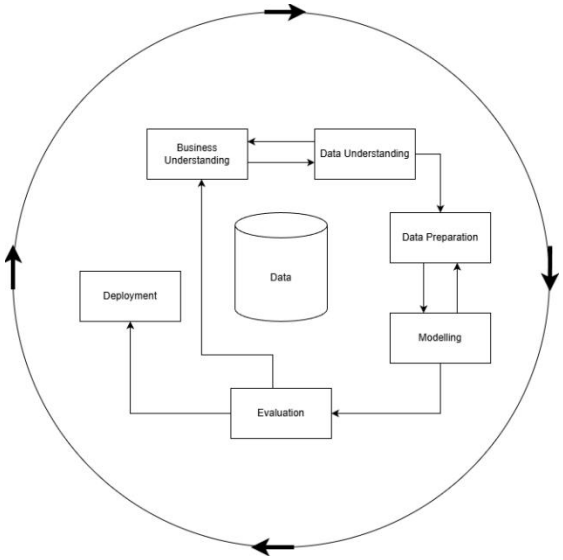
Sejumlah penelitian terdahulu telah mencoba mengeksplorasi model *multilingual* berbasis *Transformer* untuk mengatasi keterbatasan translasi pada analisis opini media sosial dan ulasan. Penelitian [9] dan penelitian [10] berhasil memanfaatkan model *multilingual* untuk klasifikasi sentimen, namun riset tersebut belum mampu mengurai polaritas sentimen berdasarkan aspek-aspek spesifik. Di sisi lain, penelitian ABSA berbasis XLM-RoBERTa umumnya masih terpaku pada analisis korpus terstruktur pada satu bahasa tertentu seperti analisis ulasan layanan medis berbahasa Uzbek [11] . Hal ini menyisakan kesenjangan besar dalam hal efektivitas model saat dihadapkan pada karakteristik data *multilingual code-mixing* ekstrem yang secara dinamis melibatkan tiga bahasa sekaligus dalam satu teks, serta kontaminasi ekspresi budaya pop atau *fandom* yang pekat di media sosial.

Untuk menjembatani kesenjangan tersebut, penelitian ini mengimplementasikan XLM-RoBERTa, yaitu model *Transformer multilingual* yang dilatih pada lebih dari 100 bahasa [12]. Model ini mampu memproses teks *multilingual* tanpa memerlukan translasi dan telah menunjukkan performa unggul dibandingkan model *multilingual* lainnya pada berbagai metrik evaluasi [13].

Penelitian ini juga memberikan kontribusi metodologis melalui komparasi performa antara prapemrosesan data yang tetap mempertahankan elemen natural khas media sosial dengan metode reduksi fitur tradisional. Eksperimen ini krusial untuk membuktikan bahwa keandalan model *Transformer* tidak hanya bergantung pada arsitekturnya, tetapi juga sangat dipengaruhi oleh bagaimana data disajikan ke dalam sistem. Dengan mengaplikasikan XLM-RoBERTa untuk ABSA pada data percakapan produk Indomie Korean Ramyeon Series di platform X, penelitian ini bertujuan untuk mengevaluasi efektivitas model dalam mengekstrak sentimen publik secara otomatis, sehingga mampu menghasilkan *insight* bisnis yang akurat dan relevan.

2. Metode Penelitian

Prosedur pelaksanaan penelitian ini dijalankan dengan mengadopsi kerangka kerja *Cross Industry Standard Process for Data Mining* (CRISP-DM) sebagai landasan metodologis utama. Kerangka kerja ini dipilih karena karakteristiknya yang sistematis, fleksibel, serta adaptif dalam menjembatani analisis data teknis dengan kebutuhan strategi bisnis [14] . Seluruh rangkaian fase pengerjaan eksperimen tersebut diilustrasikan pada Gambar 1.



Gambar 1. CRISP-DM

Metodologi penelitian ini dipetakan secara sistematis ke dalam enam tahapan utama yang saling berkesinambungan. Rangkaian proses implementasi tersebut secara kronologis diawali dengan tahap *Business Understanding*, yang kemudian dilanjutkan secara bertahap pada fase *Data Understanding* dan *Data Preparation*. Setelah data siap, alur diteruskan ke dalam fase *Modeling* dan dievaluasi kinerjanya pada tahap *Evaluation*. Seluruh rangkaian penelitian ini kemudian ditutup dengan fase *Deployment* sebagai tahap implementasi model akhir.

2.1 Business Understanding

Tahap *Business Understanding* ini didasarkan pada tantangan pengelolaan data opini terkait produk Indomie Korean Ramyeon Series di platform X. Keberagaman bahasa serta tingginya tingkat *noise* pada opini publik menjadi hambatan utama yang harus diatasi. Selama ini, studi yang melakukan ABSA untuk produk kuliner lintas budaya ini masih terbatas. Akibatnya, volume data opini yang besar di platform X tersebut belum dapat dimanfaatkan secara optimal untuk memetakan opini publik secara komprehensif.

2.2 Data Understanding

Pada tahapan *Data Understanding*, data berupa opini publik mengenai produk Indomie Korean Ramyeon Series dikumpulkan dari platform X melalui teknik *web scraping* menggunakan *tweet-harvest*. Pengambilan sampel data difokuskan pada penggunaan kata kunci (*keywords*) yang relevan dengan lini produk tersebut. Hasilnya, diperoleh dataset lintas bahasa yang mencakup tiga bahasa utama, yakni bahasa Indonesia, bahasa Inggris, dan bahasa Korea.

2.3 Data Preparation

Fase *Data Preparation* ini bertujuan untuk mentransformasikan dataset mentah menjadi format yang bersih, terstruktur, dan siap diumpankan ke dalam model *deep learning* XLM-RoBERTa. Tahapan ini memegang peranan krusial mengingat karakteristik data hasil *scraping* dari platform X cenderung korup, tidak terstruktur, dan dipenuhi *noise* yang tidak relevan [15] . Pada fase ini dilakukan proses *data cleaning, spam filtering, emoji mapping, text normalization, aspect extraction & labeling,* dan *sentiment labeling.*

Justifikasi pemilihan kategori aspek dalam penelitian ini didasarkan pada analisis frekuensi kata dominan (*term frequency analysis*) pada tahap eksplorasi awal korpus data. Untuk mengolah fitur tersebut, penelitian ini menerapkan *hybrid approach* yang mengombinasikan aturan leksikon dengan model *deep learning*. Pada tahap pertama, dokumen teks utuh terlebih dahulu diurai menjadi unit-unit klausa yang lebih kecil melalui pemotongan berbasis konjungsi kontras dan tanda baca pembatas. Selanjutnya, klasifikasi aspek pada tiap klausa dieksekusi secara presisi menggunakan pendekatan *rule-based* yang mengacu pada kamus kata kunci.

Proses pelabelan emosi awal dieksekusi secara otomatis melalui metode *automated weak labeling* memanfaatkan model multilingual *CardiffNLP Twitter-XLM-RoBERTa* untuk meningkatkan efisiensi pemrosesan data. Guna mengantisipasi adanya bias atau kesalahan prediksi oleh mesin, seluruh hasil luaran dari pelabelan otomatis tersebut kemudian melalui proses validasi manual secara komprehensif pada keseluruhan dokumen teks sebelum masuk ke tahap pemodelan akhir.

Setelah *preprocessing* selesai dilakukan akhirnya diketahui jumlah dataset per bahasa pada Tabel 1. berikut.

Tabel 1. Dataset per Bahasa setelah Preprocessing

Bahasa	Jumlah Data	Persentase
Indonesia	2.665	78.06%
Inggris	725	21.23%
Korea	24	0.70%

2.4 Modeling

Memasuki tahap *Modeling*, dataset yang telah melalui *preprocessing* dialokasikan untuk melatih arsitektur klasifikasi sentimen. Eksperimen ini mengimplementasikan model *Twitter-XLM-RoBERTa-base*, yaitu varian model *multilingual* XLM-RoBERTa yang telah melewati *pre-trained* menggunakan 198 juta data *tweet* dan dioptimalkan khusus untuk penanganan analisis sentimen. Arsitektur XLM-RoBERTa dipilih karena keunggulannya dalam merepresentasikan teks

multilingual secara simultan tanpa memerlukan pemisahan korpus bahasa [16].

2.5 Evaluation

Fase *Evaluation* dilaksanakan untuk menguji tingkat akurasi dan keandalan model akhir dalam memetakan aspek serta polaritas sentimen. Pengujian ini memanfaatkan data uji (*testing set*) yang diisolasi penuh dari proses pelatihan model. Untuk mengukur performa model secara komprehensif, metrik evaluasi yang digunakan meliputi *accuracy*, *precision*, *recall*, dan *F1-score* [17].

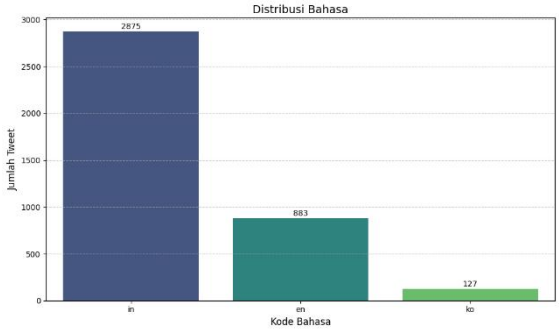
2.6 Deployment

Sebagai fase akhir, hasil ekstraksi aspek dan klasifikasi sentimen terhadap Indomie Korean Ramyeon Series diwujudkan ke dalam bentuk visualisasi data. Hal ini mencakup penyajian grafik distribusi aspek dan polaritas sentimen guna memetakan proporsi respons positif, netral, serta negatif.

3. Hasil dan Pembahasan

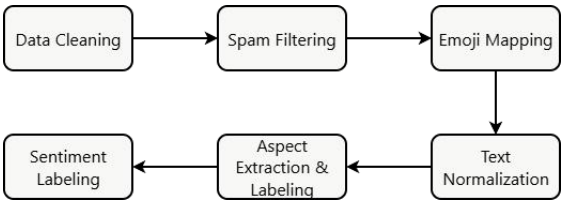
3.1 Hasil

Penelitian ini dimulai pada tahap *Business Understanding* dengan mengidentifikasi tantangan utama data opini publik pada produk Indomie Korean Ramyeon Series di platform X yang bersifat *multilingual*, penuh *code-mixing*, istilah fandom K-Pop, serta *noise* lainnya. Kombinasi yang terlalu kompleks untuk ditangani analisis sentimen konvensional. Oleh karena itu, dibutuhkan model NLP dan alur preprocessing yang tepat untuk menghasilkan input teks yang seragam dan valid sebelum dianalisis. Penelitian ini menerapkan pendekatan ABSA yang mampu mengurai sentimen publik pada setiap aspek produk secara spesifik.



Gambar 2. Distribusi Bahasa

Selanjutnya tahap *Data Understanding* seperti yang diamati pada Gambar 2, dataset mentah yang berhasil dihimpun dalam rentang waktu 31 Oktober 2024 - 30 Oktober 2025 sebanyak 3.885 data yang berasal dari tiga bahasa yang berbeda, yaitu bahasa Indonesia, bahasa Inggris, dan bahasa Korea. Selanjutnya dilakukan identifikasi awal terhadap keberadaan *noise* berupa duplikasi data, data hilang, serta teks di luar konteks produk Indomie Korean Ramyeon Series yang berpotensi mendistorsi hasil analisis.



Gambar 3. Alur Data Preparation

Berdasarkan Gambar 3, tahap *Data Preparation* diawali dengan *data cleaning*, yaitu teks dibersihkan dari URL, mention, simbol dekoratif, dan spasi berlebih. Tanda baca ekspresif seperti "!!!" dibatasi maksimal tiga karakter agar intensitas emosi tetap terjaga, sementara hashtag hanya dihapus simbolnya tanpa membuang teksnya.

Pada tahap *spam filtering, tweet* berisi promosi jualan, jastip, WTS/WTB, dan sejenisnya dieliminasi secara otomatis menggunakan daftar keyword komersial, karena konten ini bukan opini publik yang asli. Kemudian dilakukan *text normalization,* berupa normalisasi singkatan dan istilah K-Pop menggunakan kamus *lexicon*, dengan aturan pencocokan yang fleksibel terhadap kapitalisasi namun ketat terhadap batas kata agar tidak merusak kata lain yang mirip.

Selanjutnya tahap *emoji mapping*, meskipun penggunaan emoji dalam dataset ini tergolong cukup minim, proses konversi emoji ke dalam bentuk teks tetap dilakukan untuk menyeragamkan format data. Langkah ini diambil agar model dapat menangkap nuansa emosional yang terkandung dalam simbol visual tersebut secara konsisten. Melalui teknik ini, setiap emoji diubah menjadi deskripsi tekstual yang relevan dengan domain produk.

Selanjutnya pada *aspect extraction & labeling,* kategori aspek ditentukan melalui analisis frekuensi kata dan frasa yang dominan per bahasa. Hasilnya, tiga aspek utama terbentuk secara organik dari data yaitu Rasa, Kolaborasi, dan Promosi. Setiap *tweet* kemudian dipotong menjadi klausa-klausa lebih kecil menggunakan pemisah kontras ("tetapi", "namun") dan tanda baca (;), lalu masing-masing klausa dilabeli aspeknya dengan pendekatan *rule-based* berdasarkan kamus kata kunci yang telah ditentukan. *Tweet* yang tidak termasuk salah satu dari tiga aspek akan dikategorikan ke dalam kelas *Others* untuk kemudian dieliminasi sebelum memasuki fase. *Modeling*.

Setelah itu *sentiment labeling*, yaitu dengan pelabelan sentimen awal dilakukan secara otomatis menggunakan model *Twitter-XLM-RoBERTa-base* untuk efisiensi. Hasilnya kemudian divalidasi

secara manual secara keseluruhan. Setelah melalui *preprocessing*, akhirnya data yang siap untuk tahap selanjutnya sebanyak 3.414 data.

Sebelum masuk ke *Modeling*, dataset diformat terlebih dahulu menggunakan struktur *Aspect-Context Pair* (ACP) dengan token pemisah [SEP] agar model bisa memahami relasi antara aspek dan klausanya. Label sentimen dikonversi ke numerik (0, 1, 2), lalu dataset dibagi dengan *Stratified Split* menjadi 80% data latih, 10% validasi, dan 10% uji. Tokenisasi dilakukan menggunakan *SentencePiece tokenizer* dari arsitektur XLM-RoBERTa dengan batas 160 token yang disesuaikan dengan panjang *tweet* di platform X.

Pelatihan model dikembangkan menggunakan *CustomTrainer* dengan penanganan bobot kelas (*class weights*) untuk memitigasi ketidakseimbangan distribusi kelas (*class imbalance*). Konfigurasi tersebut disesuaikan dengan parameter optimal untuk menjaga kestabilan pelatihan sekaligus mencegah *overfitting*. Sistem juga dilengkapi fitur *early stopping* yang otomatis menghentikan iterasi ketika performa model stagnan, lalu mengunci versi terbaiknya.

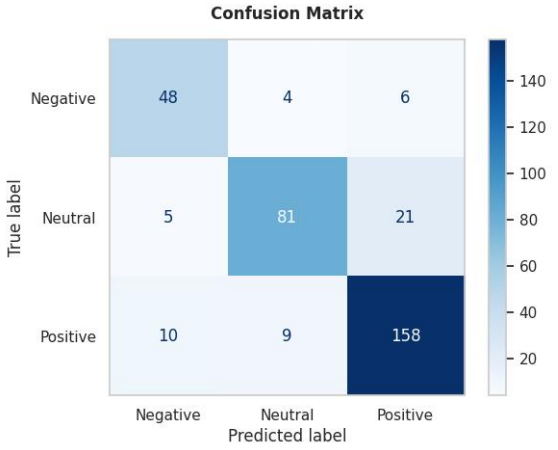
Hasil performa model menunjukkan perkembangan optimal di awal proses, mencapai puncaknya pada *epoch* ke-2 dengan tingkat akurasi sebesar 85,34% dan *F1-score* sebesar 84,84%. Setelah iterasi tersebut, nilai *validation loss* mengalami peningkatan dari 0,35 menjadi 0,61 pada epoch ketiga dan terus melonjak hingga 0,94 pada epoch keempat yang menandakan adanya sinyal *overfitting*. Oleh karena itu, mekanisme *early stopping* secara otomatis menghentikan proses pelatihan pada *epoch* ke-4 dan mengunci bobot model terbaik dari epoch ke-2 sebagai model final.

Tabel 2. Evaluation Matrix

	Precision	Recall	F1-Score	Support
Negative	0.7619	0.8276	0.7934	58
Neutral	0.8617	0.7570	0.8060	107
Positive	0.8541	0.8927	0.8729	177
Accuracy	-	-	0.8392	342
Macro Avg	0.8259	0.8258	0.8241	342
Weighted Avg	0.8408	0.8392	0.8385	342

Berdasarkan Tabel 2, model memperoleh nilai *accuracy* keseluruhan sebesar 83,92%. Ditinjau per kelas sentimen, kelas *Positive* mencatatkan *F1-score* tertinggi sebesar 87,29% dengan *recall* mencapai 89,27%, yang mengindikasikan bahwa model memiliki kemampuan tinggi dalam mengenali sentimen positif dari keseluruhan data aktual yang tersedia. Kelas *Neutral* memperoleh *F1-score* sebesar 80,60% dengan *precision* 86,17% dan *recall* 75,70%. Sementara itu, kelas *Negative* memperoleh *F1-score* sebesar 79,34%, dengan *precision* 76,19% dan *recall* yang cukup tinggi sebesar 82,76%. Tingginya nilai *recall* pada kelas minoritas ini menunjukkan bahwa model mampu menjaring teks bernada negatif secara lebih sensitif dan tidak lagi cenderung melewatkannya.

Secara agregat, nilai *Macro Average F1-score* diperoleh sebesar 82,41%, sedangkan *Weighed Average F1-score* sebesar 83,85%. Selisih yang sangat tipis antara nilai *Macro* dan *Weighted Average* ini mengindikasikan bahwa dampak negatif dari *class imbalance* pada data uji telah berhasil dimitigasi dengan optimal melalui fungsi *class weights*. Meskipun secara faktual kelas *Positive* tetap mendominasi dengan 177 data dibandingkan kelas *Negative* yang hanya berjumlah 58 data, model terbukti tetap mampu memberikan penilaian performa yang adil dan merata di seluruh kategori sentimen dengan total data uji keseluruhan mencapai 342 data.



Gambar 4. Confusion Matrix

Berdasarkan Gambar 4, visualisasi *confusion matrix* menunjukkan akurasi dominan pada diagonal utama, dengan akurasi prediksi tertinggi pada kelas *Positive* sebesar 89,27%, diikuti kelas *Negative* sebesar 82,76%, dan kelas *Neutral* sebesar 75,70%. Kesalahan klasifikasi paling masif bergeser pada 21 data *Neutral* (19,63%) yang keliru teridentifikasi sebagai *Positive*. Anomali ini dipicu oleh adanya bias kontekstual media sosial, di mana model mengaitkan *attention score* nama *brand ambassador* "NewJeans" secara pekat ke polaritas positif, sehingga menggeser batas keputusan (*decision boundary*) dari teks netral/informasional ke arah positif.

Pola misklasifikasi terbesar kedua terjadi pada 10 data *Positive* (5,65%) yang meleset menjadi *Negative*, terindikasi akibat keterbatasan model dalam memilah struktur linguistik informal yang kompleks seperti sarkasme atau kosakata hiperbola bernada ekstrem. Meskipun demikian, rendahnya angka kesalahan tebakan data *Negative* yang

meleset ke arah *Positive* (10,34%) memberikan justifikasi ilmiah yang kuat. Fenomena penekanan *positive bias* ini membuktikan bahwa implementasi *Custom Trainer* berbasis *Class Weights* berhasil memitigasi dampak buruk ketidakseimbangan distribusi data uji (*class imbalance*), sehingga memaksa model bertindak sensitif dan hati-hati sebelum melabeli sentimen minoritas.

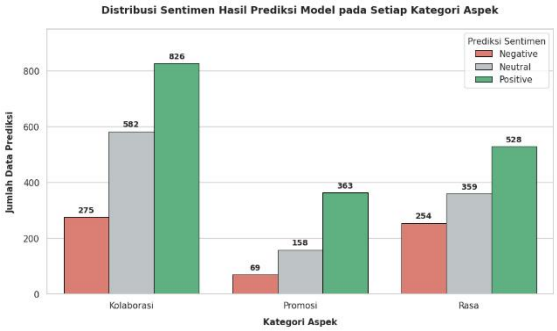
Selain itu, dilakukan evaluasi untuk membandingkan efektivitas dua skema *preprocessing* terhadap performa klasifikasi sentimen model XLM-RoBERTa. Dengan menggunakan data uji yang sama, penelitian ini menguji sejauh mana model mampu mempertahankan akurasi ketika fitur teks diproses melalui pendekatan khusus atau alur yg digunakan di awal (Alur A) dibandingkan dengan pendekatan reduksi fitur tradisional (Alur B). Alur B melakukan penyederhanaan teks secara agresif dengan melakukan *case folding*, penghapusan total emoji serta tanda baca, serta menerapkan *stopword removal* dan *stemming* untuk mengubah kata menjadi bentuk dasarnya.

Tabel 2. Perbandingan Alur Preprocessing

Alur Preprocessing	Accuracy	Macro Avg F1-Score	Weighted Avg F1-Score	Validation Loss
Alur A: Khusus	83.92%	82.41%	83.85%	0.3920
Alur B: Klasik	78.95%	75.93%	78.84%	0.6974
Selisih	+4.97%	+6.48%	+5.01%	-0.3054

Alur A terbukti lebih unggul karena mampu menjaga konteks asli teks, mulai dari penggunaan emoji dan bahasa gaul hingga pemisahan klausa yang lebih rapi. Hal ini membuat model XLM-RoBERTa bekerja jauh lebih maksimal dibandingkan dengan metode klasik pada Alur B yang justru sering merusak struktur kalimat alami lewat proses *stemming* dan penghapusan kata tugas secara berlebihan. Dengan mempertahankan fitur-fitur kontekstual tersebut, model jadi lebih cerdas dalam menangkap emosi di balik teks, sehingga tebakannya pun menjadi lebih akurat.

Setelah model dievaluasi, model digunakan untuk memprediksi sentimen pada keseluruhan data yang telah dikategorikan berdasarkan tiga aspek, yaitu Kolaborasi, Promosi, dan Rasa.



Gambar 5. Hasil Sentimen per Aspek

Berdasarkan Gambar 5, data hasil prediksi terdistribusi ke dalam tiga kategori aspek, di mana aspek Kolaborasi mencatatkan volume terbesar dengan total 1.683 data (826 *Positive*, 582 *Neutral*, 275 *Negative*), diikuti aspek Rasa sebanyak 1.141 data (528 *Positive*, 359 *Neutral*, 254 *Negative*), dan aspek Promosi sebagai klaster terkecil dengan 590 data (363 *Positive*, 158 *Neutral*, 69 *Negative*). Pada seluruh kategori tersebut, sebaran data secara konsisten didominasi oleh prediksi sentimen *Positive*, yang kemudian diikuti oleh polaritas *Neutral* dan *Negative*. Selain itu, aspek Kolaborasi menjadi kategori dengan proporsi sentimen *Neutral* tertinggi secara absolut (582 data), sementara aspek Promosi mencatatkan jumlah sentimen *Negative* paling rendah yaitu sebanyak 69 data.

3.2 Pembahasan

Penggunaan model *Twitter-XLM-RoBERTa-base* pada tugas ABSA untuk produk Indomie Korean Ramyeon Series terbukti memberikan performa generalisasi yang sangat baik, terutama saat dihadapkan pada opini publik di platform X. Keunggulan ini utamanya didorong oleh adanya *domain alignment* yang kuat. Mengingat dataset penelitian ini didominasi oleh bahasa informal, *slang*, serta fenomena *code-switching* yang masif khas media sosial, model ini mampu bekerja secara optimal karena telah dilatih sebelumnya menggunakan jutaan *tweet* asli [18]. Dibandingkan dengan model bahasa standar lainnya, *Twitter-XLM-RoBERTa* memiliki representasi vektor kontekstual yang jauh lebih matang dalam mengenali struktur bahasa tidak baku yang sering ditemui di platform tersebut.

Karakteristik *multilingual* yang informal dalam dataset ini memberikan keunggulan bagi *Twitter-XLM-RoBERTa* dibandingkan model *baseline* lainnya. Model *monolingual* seperti *IndoBERT*, cenderung mengalami *sub-word shattering* saat menemui entitas asing atau *slang* internet, yang mengakibatkan hilangnya konteks semantik secara utuh [19]. Di sisi lain, mBERT kurang adaptif terhadap gaya bahasa kasual media

sosial karena pelatihannya berbasis korpus formal Wikipedia [20] . Sebaliknya, arsitektur Twitter-XLM-RoBERTa memanfaatkan tokenisasi *SentencePiece* dan *Masked Language Modeling* (MLM) lintas bahasa yang terbukti jauh lebih kokoh dalam mempertahankan representasi konteks semantik kalimat, sekalipun dalam struktur *code-mixing* [18].

Tantangan utama dalam analisis sentimen media sosial sering kali muncul dari ketidakseimbangan distribusi data (*class imbalance*), di mana model cenderung mengabaikan kelas minoritas [21]. Untuk mengatasi masalah ini, penelitian ini menerapkan *CustomTrainer* dengan mekanisme pembobotan kelas (*class weight adjustment*). Strategi ini terbukti efektif menjaga sensitivitas model dalam mendeteksi sentimen negatif secara adil, selaras dengan temuan pada studi sebelumnya [22]. Hal ini menjadi sangat krusial, mengingat keluhan konsumen atau sentimen negatif sering kali menyimpan *insight* paling berharga untuk evaluasi kualitas produk ke depan.

Aspek Kolaborasi mencatatkan volume data terbesar yang menegaskan bahwa strategi kolaborasi Indomie dengan *girlgroup* NewJeans merupakan *attention driver* utama di platform X. Tingginya proporsi sentimen *Neutral* pada aspek ini merepresentasikan tingginya aktivitas diskusi publik yang bersifat interogatif maupun pernyataan objektif tanpa polaritas emosi, seperti pertanyaan mengenai ketersediaan bonus *photocard* atau *postcard* NewJeans hingga sekadar *tweet* informatif mengenai kolaborasi antara Indomie dengan NewJeans. Di sisi lain, munculnya sentimen negatif sebesar 16,34% didorong oleh faktor eksternal di luar produk, yaitu adanya kekhawatiran dan sentimen negatif publik yang mengaitkan kolaborasi ini dengan konflik korporasi eksternal yang sedang terjadi antara agensi ADOR dan pihak manajemen HYBE. Sebagian pengguna memproyeksikan konflik industri musik K-Pop tersebut ke dalam *tweet* mereka, sehingga memengaruhi atmosfer perbincangan mengenai *brand ambassador* NewJeans pada lini produk Indomie ini.

Aspek Rasa menempati posisi kedua dengan tingkat sentimen positif yang kuat, menandakan bahwa profil rasa yang ditawarkan berhasil memenuhi ekspektasi umum konsumen Indonesia terhadap cita rasa kuliner Korea. Kendati demikian, aspek ini juga mencatatkan persentase sentimen negatif tertinggi (22,26%). Melalui analisis mendalam terhadap klaster data, keluhan tersebut karena sekadar tidak menyukai varian rasanya atau konsumen yang mengeluhkan tingkat kepedasan yang terlalu tinggi sehingga tidak sesuai dengan toleransi konsumsi mereka.

Aspek Promosi memiliki volume paling rendah, tetapi aspek ini menorehkan tingkat kepuasan tertinggi secara persentase dengan angka sentimen positif mencapai 61,53% dan sentimen negatif minimum. Karakteristik ini menunjukkan bahwa eksekusi kampanye pemasaran kreatif di lapangan, seperti instalasi *billboard* dan materi iklan visual dinilai sangat estetik dan sukses menyentuh sisi emosional audiens target maupun komunitas penggemar. Keberhasilan promosi visual ini secara efektif mampu meredam potensi munculnya narasi penolakan publik terhadap aktivitas periklanan merek.

4. Kesimpulan

Penelitian ini membuktikan efektivitas implementasi XLM-RoBERTa dalam menjawab tantangan *Aspect-Based Sentiment Analysis* (ABSA) pada data multilingual yang sarat fenomena *code-switching* di platform X. Secara teoretis, hasil riset ini mengonfirmasi bahwa penyelarasan domain (*domain alignment*) antara data pra-pelatihan dengan karakteristik dataset target jauh lebih krusial dibandingkan pendekatan translasi teks formal. Keandalan arsitektur Transformer terbukti lewat mekanisme *self-attention* yang mampu mempertahankan nuansa semantik asli dari tiga bahasa secara simultan, sekaligus mengukuhkan bahwa strategi *data preparation* dengan mempertahankan emoji, *slang*, dan struktur klausa memberikan performa yang lebih optimal dibandingkan metode reduksi fitur tradisional. Selain itu, penggunaan *custom class weights* terbukti menjadi solusi efektif dalam memitigasi bias prediksi akibat ketidakseimbangan data (*class imbalance*).

Secara praktis, model ABSA yang dikembangkan dalam penelitian ini berkontribusi sebagai instrumen *data-driven* otomatis untuk mendukung pengambilan keputusan bisnis yang cepat dan akurat bagi manajemen perusahaan. Melalui pemetaan sentimen per aspek, pihak manajemen dapat mengekstrak *actionable insights* yang objektif untuk mengevaluasi dampak strategis kolaborasi merek dengan *brand ambassador*, membaca pergeseran preferensi pasar, serta memitigasi krisis reputasi di ruang digital. Implementasi teknologi AI ini berhasil mentransformasi data percakapan media sosial yang tidak terstruktur menjadi metrik evaluasi pemasaran yang bernilai guna tinggi.

Meskipun demikian, penelitian ini masih dihadapkan pada sejumlah limitasi, antara lain tantangan model dalam mendeteksi sentimen negatif atau netral yang disampaikan secara implisit dan sarkastik, pendekatan ekstraksi aspek yang

masih bersifat *rule-based* manual sehingga berpotensi melewatkan aspek laten organik, serta ketimpangan distribusi bahasa yang didominasi oleh bahasa Indonesia informal. Merespons keterbatasan tersebut, penelitian mendatang disarankan untuk mengadopsi pendekatan ekstraksi aspek dinamis seperti *end-to-end* ABSA yang mampu mengidentifikasi aspek secara otomatis. Selain itu, optimasi penanganan *class imbalance* dapat ditingkatkan ke skala yang lebih *advanced* menggunakan fungsi *Focal Loss* disertai pengayaan proporsi korpus bahasa minoritas agar kemampuan *multilingual* model dapat dieksploitasi secara maksimal.

Daftar Pustaka

[1] Statista, "Number of social media users worldwide from 2017 to 2030." [Online]. Available: https://www.statista.com/forecasts/278414/number-of-worldwide-social-network-users/

[2] H. Muscolino, A. Machado, J. Rydning, and D. Vesset, *Untapped Value: What Every Executive Needs to Know About Unstructured Data*. Needham, MA, USA: International Data Corporation (IDC), 2023.

[3] O. Toffaha and L. Tashtoush, "The Impact of Artificial Intelligence on Marketing Strategies and Business Sustainability," *Journal for Social Science Archives*, vol. 1, no. 2, pp. 80–90, 2023, doi: 10.59075/jssa.v1i2.10.

[4] B. Liu, *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions*. Cambridge University Press, 2015.

[5] A. P. Al Aufar and A. Romadhony, "Aspect-based Sentiment Analysis on Beauty Product Reviews using BERT and Long Short-Term Memory," *Jurnal Nasional Pendidikan Teknik Informatika (JANAPATI)*, vol. 14, no. 2, pp. 364–373, 2025, doi: 10.23887/janapati.v14i2.94392.

[6] Z. N. Wandi, P. Muljono, and F. Fuad Cholagi, "Penerapan Retorika Visual dalam Iklan Indomie Korean Ramyeon Series X NewJeans di Media Youtube," *Jurnal Komputer, Informasi dan Teknologi*, vol. 5, no. 1, p. 18, 2025, doi: 10.53697/jkomitek.v5i1.2468.

[7] W. Wongso, D. S. Setiawan, S. Limcorn, and A. Joyoadikusumo, "NusaBERT: Teaching IndoBERT to be Multilingual and Multicultural," in *Proceedings of the Second Workshop in South East Asian Language Processing*, Association for Computational Linguistics, 2025, pp. 10–26. Accessed: Jun. 23, 2026. [Online]. Available: https://aclanthology.org/2025.sealp-1.2/

[8] G. D. Rahma, "Cross-Language Sentiment Analysis in Multilingual Social Media Discourse," *International Journal of Research and Applied Technology (INJURATECH)*, vol. 4, no. 2, pp. 366–369, 2024, Accessed: Jun. 23, 2026. [Online]. Available: https://ojs.unikom.ac.id/index.php/injuratech/article/view/19670

S. S. Almalki, "Sentiment Analysis and Emotion Detection Using Transformer Models in Multilingual Social Media Data," *International Journal of Advanced Computer Science and Applications*, vol. 16, no. 3, 2025, doi: 10.14569/IJACSA.2025.0160332.

] N. N. I. Prova, V. Ravi, M. P. Singh, V. K. Srivastava, S. Chippagiri, and A. P. Singh, "Multilingual sentiment analysis in e-commerce customer reviews using GPT and deep learning-based weighted-ensemble model," *International Journal of Cognitive Computing in Engineering*, vol. 7, no. 1, pp. 268–286, Dec. 2026, doi: 10.1016/j.ijcce.2025.10.003.

[11] B. Elov and R. Alaev, "Aspect-Based Sentiment Analysis of Hospital Service Reviews Using BERTbek and XLM-R," *SSRN*, 2026, [Online]. Available: https://ssrn.com/abstract=6570837

[12] H. Sanjaya *et al.*, "Forest Fire Location and Time Recognition in Social Media Text Using Xlm-Roberta," *JITK (Jurnal Ilmu Pengetahuan dan Teknologi Komputer)*, vol. 10, no. 4, pp. 749–758, 2025, doi: 10.33480/jitk.v10i4.6194.

[13] M. M. A. Paramarta, R. Dwiyansaputra, and R. P. Rassy, "Performance Analysis of Multilingual and Monolingual Models in Predicting Indonesian," *Jurnal Teknologi Informasi, Komputer dan Aplikasinya (JTIKA)*, vol. 7, no. 2, pp. 237–246, 2025, Accessed: Jun. 23, 2026. [Online]. Available: https://doi.org/10.29303/jtika.v7i2.482

[14] D. T. Larose, *Discovering Knowledge in Data*. Hoboken, New Jersey: John Wiley & Sons, Inc., 2005.

[15] J. S. Saltz, "CRISP-DM for Data Science: Strengths, Weaknesses and Potential Next Steps," *2021 IEEE International Conference on Big Data (Big Data)*, pp. 2337–2344, 2021, doi: 10.1109/BigData52589.2021.9671634.

[16] A. R. Subarkah and R. Imanda, "Penerapan XLM-R dan TD-IF+Consiner Similarity untuk Deteksi Berita Hoax Multilanguage pada Aplikasi Mobile Berbasis Natural Language Processing," *METIK Jurnal*, vol. 9, pp. 215–224, 2025, doi: 10.47002/metik.v9i1.1066.

[17] J. Erbani, P.-édouard Portier, E. Egyed-zsigmond, and D. Nurbakova, "Confusion Matrices : A Unified Theory," *IEEE Access*, vol. 12, pp. 181372–181419, 2024, doi: 10.1109/ACCESS.2024.3507199.

[18] F. Barbieri, L. E. Anke, and J. Camacho-collados, "XLM-T : Multilingual Language Models in Twitter for Sentiment Analysis and Beyond," in *Proceedings of the Thirteenth Language Resources and Evaluation Conference*,

European Language Resources Association, 2022, pp. 258–266. Accessed: Jun. 23, 2026. [Online]. Available: https://aclanthology.org/2022.lrec-1.27/

[19] F. Koto, J. H. Lau, and T. Baldwin, "INDOBERTWEET: A Pretrained Language Model for Indonesian Twitter with Effective Domain-Specific Vocabulary Initialization," in *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, Association for Computational Linguistics, 2021, pp. 10660–10668. Accessed: Jun. 23, 2026. [Online]. Available: https://aclanthology.org/2021.emnlp-main.833/

[20] G. I. Winata, S. Cahyawijaya, Z. Liu, Z. Lin, A. Madotto, and P. Fung, "Are Multilingual Models Effective in Code-Switching?," in *Proceedings of the Fifth Workshop on Computational Approaches to Linguistic Code-Switching*, Association for Computational Linguistics, 2021, pp. 200–208. doi: 10.26615/978-954-452-056-4_020.

[21] N. Lin, M. Zeng, X. Liao, W. Liu, A. Yang, and D. Zhou, "Addressing class-imbalance challenges in cross-lingual aspect-based sentiment analysis : Dynamic weighted loss and anti-decoupling," *Expert Syst. Appl.*, vol. 257, 2024, Accessed: Jun. 23, 2026. [Online]. Available: https://doi.org/10.1016/j.eswa.2024.125059

[22] A. P. Maretta and A. Meiriza, "Aspect-Based Sentiment Analysis of Hospital Service Reviews Using Fine-Tuned IndoBERT," *Journal of Applied Informatics and Computing (JAIC)*, vol. 9, no. 5, 2025, Accessed: Jun. 23, 2026. [Online]. Available: https://doi.org/10.30871/jaic.v9i5.10765