

Published online on the page: https://journal.upp.ac.id/index.php/rjti/index

R J T I

Riau Jurnal Teknik Informatika

| E - ISSN (*Online*) 2829-2529 |

Analisis Sentimen Berbasis Aspek Ulasan GoFood di X dengan Ekstraksi Aspek *Rule-Based* dan Klasifikasi IndoBERT

Nazlah Ayunda Syahputri Jacub[1,*], Agustinus Jacobus[2], Muhamad Dwisnanto Putro[3]

[1,2,3] Universitas Sam Ratulangi, Manado, Indonesia

Informasi Artikel

Sejarah Artikel:
Submit: 27 Juli 2026
Revisi: 10 Oktober 2026
Diterima: 22 Oktober 2026
Diterbitkan: 02 November 2026

Kata Kunci

Analisis Sentimen Berdasarkan Aspek, GoFood, IndoBERT, Rule-Based, Layanan Pesan-Antar Makanan, Platform X.

Korespondensi

E-mail:
nazlahjacub026@student.unsrat.ac.id

A B S T R A C T

Social media reviews provide insights into user perceptions of food delivery services, but general sentiment analysis does not identify specific service aspects. This study applies Aspect-Based Sentiment Analysis (ABSA) using a hybrid Rule-Based and IndoBERT approach to GoFood reviews on X. Rule-Based methods extract aspects using keywords and structured rules, while IndoBERT classifies sentiment based on textual context. From 28,763 Indonesian-language tweets obtained from Kaggle, 5,657 aspect–sentiment pairs were obtained after preprocessing, aspect extraction, auditing, and manual verification. The dataset covers four aspects: price, service, food quality, and application. Using a StratifiedGroupKFold split of 70:15:15, IndoBERT achieved 90.31% accuracy and a Macro F1-score of 0.8883 on 846 test samples. Food quality achieved the highest aspect-specific accuracy (92.62%), followed by price (90.40%), service (90.00%), and application (87.88%). Price was the most frequently identified aspect, while positive sentiment dominated food quality (60.4%) and neutral sentiment dominated the other three aspects. These findings provide detailed aspect-level insights into user perceptions and can help identify service aspects to maintain or examine further. However, the impact on service quality improvement has not been directly evaluated.

Abstrak

Ulasan media sosial dapat menggambarkan persepsi pengguna terhadap layanan pesan-antar makanan, tetapi analisis sentimen umum belum mengidentifikasi aspek layanan yang menjadi sasaran sentimen. Penelitian ini menerapkan Aspect-Based Sentiment Analysis (ABSA) dengan menggabungkan Rule-Based untuk mengekstraksi aspek berdasarkan kata kunci dan aturan terstruktur serta IndoBERT untuk mengklasifikasikan sentimen berdasarkan konteks teks. Kombinasi ini memungkinkan analisis sentimen dilakukan secara lebih terperinci pada setiap aspek layanan GoFood di platform X. Rule-Based digunakan untuk mengekstraksi aspek berdasarkan kata kunci dan aturan terstruktur, sedangkan IndoBERT mengklasifikasikan sentimen berdasarkan konteks teks. Dari 28.763 tweet berbahasa Indonesia yang diperoleh dari Kaggle, dihasilkan 5.657 pasangan aspek–sentimen yang mencakup aspek harga, pelayanan, kualitas makanan, dan aplikasi. Dengan pembagian data StratifiedGroupKFold sebesar 70:15:15, IndoBERT memperoleh akurasi 90,31% dan Macro F1-score 0,8883 pada 846 data uji. Akurasi tertinggi diperoleh pada aspek kualitas makanan (92,62%), diikuti harga (90,40%), pelayanan (90,00%), dan aplikasi (87,88%). Harga merupakan aspek yang paling banyak ditemukan, sedangkan sentimen positif mendominasi kualitas makanan (60,4%) dan sentimen netral mendominasi tiga aspek lainnya. Temuan ini memberikan informasi sentimen yang lebih terperinci untuk memahami persepsi pengguna dan mengidentifikasi aspek layanan

Creative Commons Attribution-ShareAlike 4.0 International License. *Some rights reserved*

yang perlu dipertahankan atau ditinjau lebih lanjut. Dampaknya terhadap peningkatan kualitas layanan belum dievaluasi secara langsung.

This is an open access article under the CC–BY-SA license

1. Pendahuluan

Perkembangan layanan pesan-antar makanan digital mendorong pengguna untuk menyampaikan pengalaman dan penilaian terhadap layanan melalui media sosial. Ulasan tersebut dapat menjadi sumber informasi untuk memahami persepsi pengguna terhadap kualitas makanan, harga, pelayanan, dan aplikasi [1], [2]. Namun, banyaknya data serta penggunaan bahasa informal pada media sosial membuat analisis secara manual kurang efisien [3], [4]. Oleh karena itu, diperlukan pendekatan analisis sentimen yang mampu mengolah ulasan secara otomatis untuk memperoleh informasi yang relevan bagi evaluasi layanan.

Analisis sentimen konvensional umumnya mengklasifikasikan opini pada tingkat teks secara keseluruhan sehingga belum tentu dapat menunjukkan penilaian pengguna terhadap setiap aspek layanan. Satu ulasan dapat memuat sentimen berbeda, misalnya positif terhadap kualitas makanan tetapi negatif terhadap pelayanan. *Aspect-Based Sentiment Analysis* (ABSA) mengatasi keterbatasan tersebut dengan mengidentifikasi aspek yang dibahas dan menentukan sentimen pada setiap aspek secara terpisah [5], [6]. Dalam konteks GoFood, pendekatan ini memungkinkan opini pengguna dipetakan berdasarkan aspek kualitas makanan, harga, pelayanan, dan aplikasi. Informasi tersebut memberikan gambaran yang lebih terperinci mengenai aspek yang memperoleh tanggapan positif, netral, atau negatif, sehingga dapat menjadi bahan pendukung evaluasi layanan [7], [8].

Penelitian ABSA telah dikembangkan pada berbagai domain dengan pendekatan yang beragam. Yusni Adhelina et al. menerapkan model CNN dan LSTM pada ulasan *e-commerce* [9], sedangkan Wafda et al. menggunakan IndoBERT untuk ABSA pada data Twitter mengenai Kurikulum Merdeka [10]. Rayie et al. mengkaji sentimen berdasarkan dimensi kualitas produk menggunakan IndoBERT dan RoBERTa [11]. Penelitian lain juga mengeksplorasi penerapan ABSA pada umpan balik mahasiswa dan ulasan layanan digital [12], [13]. Hasil penelitian tersebut menunjukkan bahwa ABSA dapat diterapkan pada berbagai domain, tetapi karakteristik data, metode ekstraksi aspek, dan tujuan klasifikasi perlu dipertimbangkan ketika pendekatan tersebut diterapkan pada ulasan layanan pesan-antar makanan.

IndoBERT dipilih sebagai model klasifikasi sentimen karena dirancang untuk memproses teks berbahasa Indonesia dengan representasi kontekstual berbasis *Transformer* [14], [15]. Kemampuan ini relevan untuk mengolah ulasan media sosial yang mengandung variasi kosakata, ungkapan informal, dan konteks kalimat [16]. Sementara itu, pendekatan *Rule-Based* digunakan untuk mengekstraksi aspek berdasarkan kamus kata kunci dan aturan yang ditentukan secara eksplisit. Pendekatan ini dipilih karena hasil ekstraksinya dapat ditelusuri berdasarkan aturan yang digunakan dan tidak terlalu bergantung pada ketersediaan data berlabel untuk melatih model ekstraksi aspek [17], [18]. Meskipun demikian, keberhasilannya bergantung pada kelengkapan kamus dan aturan dalam menangani variasi bahasa. Oleh karena itu, penggabungan *Rule-Based* dan IndoBERT memungkinkan ekstraksi aspek dilakukan secara terstruktur, sedangkan klasifikasi sentimen dilakukan dengan mempertimbangkan konteks teks.

Sejumlah penelitian terdahulu telah mengkaji penggunaan model Transformer dan pendekatan hibrida dalam analisis sentimen. Fatihaturrahmah dan Ditha Tania menggabungkan *Rule-Based* dan IndoBERT untuk menganalisis persepsi keamanan pada ulasan aplikasi DANA [19]. Namun,

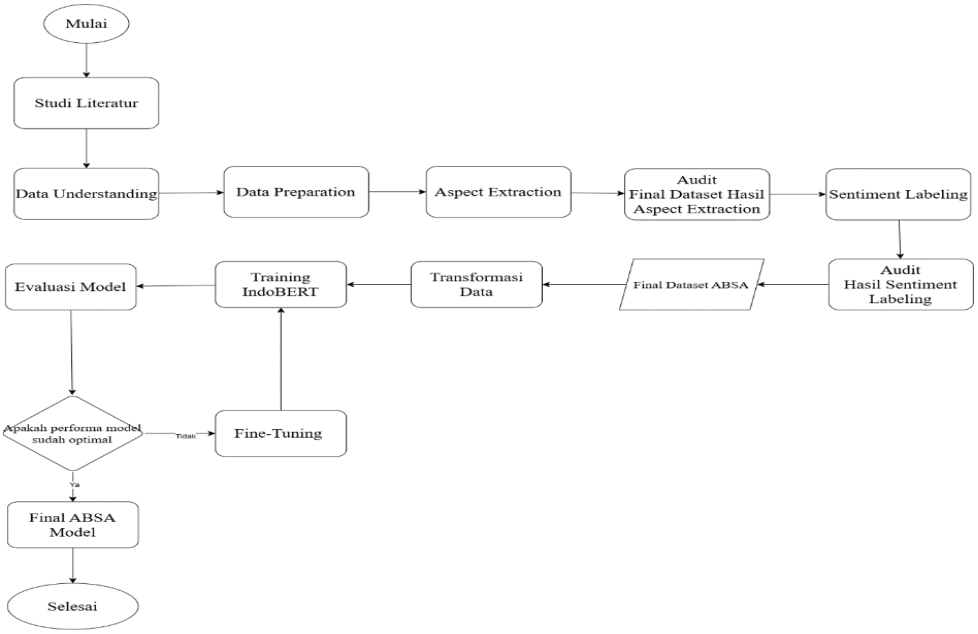
penelitian tersebut menggunakan aspek keamanan yang telah ditentukan sebelumnya sehingga berbeda dari pendekatan yang melakukan ekstraksi beberapa aspek dari teks ulasan. Sementara itu, Lin dan Nuha menerapkan pendekatan *hybrid deep learning* pada data media sosial, tetapi berfokus pada analisis sentimen tingkat dokumen [20]. Penelitian Setiawan et al. juga menggunakan IndoBERT untuk klasifikasi sentimen pada tingkat dokumen [13]. Perbedaan tersebut menunjukkan adanya peluang untuk menerapkan pendekatan yang secara eksplisit mengekstraksi beberapa aspek dari ulasan layanan pesan-antar makanan, kemudian mengklasifikasikan sentimen pada setiap aspek. Selain itu, penggunaan pendekatan hibrida pada domain GoFood di platform X perlu dikaji dengan mempertimbangkan karakteristik bahasa dan konteks ulasan yang digunakan.

Berdasarkan permasalahan tersebut, penelitian ini menerapkan pendekatan hibrida *Rule-Based* dan IndoBERT untuk melakukan ABSA pada ulasan pengguna GoFood di platform X. *Rule-Based* digunakan untuk mengekstraksi empat aspek layanan, yaitu kualitas makanan, harga, pelayanan, dan aplikasi, sedangkan IndoBERT digunakan untuk mengklasifikasikan sentimen positif, netral, dan negatif pada setiap aspek yang teridentifikasi. Data dipersiapkan melalui preprocessing, pelabelan sentimen awal menggunakan RoBERTa, serta audit dan verifikasi manual sebelum digunakan dalam pelatihan dan evaluasi model. Kontribusi penelitian ini terletak pada penerapan ekstraksi beberapa aspek secara eksplisit yang diikuti klasifikasi sentimen berbasis konteks pada domain ulasan GoFood di X. Hasil penelitian diharapkan dapat memberikan gambaran sentimen yang lebih terperinci berdasarkan aspek layanan dan menjadi informasi pendukung dalam evaluasi pengalaman pengguna.

2. Metode Penelitian

2.1. Tahapan Penelitian

Penelitian ini menggunakan pendekatan hibrida *Rule-Based* dan IndoBERT untuk menerapkan *Aspect-Based Sentiment Analysis* (ABSA) pada ulasan pengguna GoFood di platform X. Tahapan penelitian mengacu pada kerangka *Cross-Industry Standard Process for Data Mining* (CRISP-DM) yang disesuaikan dengan kebutuhan penelitian. Proses dimulai dari pemahaman data, persiapan data, ekstraksi aspek, audit dan koreksi manual, pelabelan sentimen awal menggunakan RoBERTa, serta verifikasi label. Dataset yang telah diverifikasi selanjutnya digunakan untuk melatih dan mengevaluasi IndoBERT. Alur penelitian secara keseluruhan ditunjukkan pada Gambar 1.



Gambar 1. Tahapan Penelitian

2.2. Data Understanding

Dataset penelitian berupa 28.763 ulasan pengguna GoFood berbahasa Indonesia yang diperoleh dari Kaggle. Data memuat teks ulasan yang membahas pengalaman pengguna terhadap layanan GoFood. Penelitian berfokus pada empat aspek, yaitu kualitas makanan, harga, pelayanan, dan aplikasi, dengan tiga kelas sentimen: positif, netral, dan negatif.

Pada tahap awal, sentimen global ulasan diprediksi menggunakan model *w11wo/indonesian-roberta-base-sentiment-classifier*. Hasil prediksi awal terdiri atas 9.131 ulasan positif, 9.350 ulasan netral, dan 10.282 ulasan negatif. Label tersebut digunakan sebagai bantuan pelabelan awal dan tidak langsung dianggap sebagai label sentimen final pada setiap aspek. Hal ini karena sentimen suatu ulasan secara keseluruhan belum tentu mencerminkan sentimen terhadap aspek tertentu.

Data dari platform X memiliki karakteristik bahasa informal, seperti singkatan, slang, variasi ejaan, pengulangan karakter, dan ungkapan yang bergantung pada konteks. Karakteristik tersebut dipertimbangkan dalam preprocessing, ekstraksi aspek, dan verifikasi label.

2.3. Data Preparation

Tahap persiapan data dilakukan untuk meningkatkan kebersihan dan relevansi teks sebelum ekstraksi aspek. Dari 28.763 ulasan awal, sebanyak 8 data kosong dan 1.890 data duplikat dihapus sehingga tersisa 26.865 ulasan. Selanjutnya, penyaringan *noise* menghasilkan 17.220 ulasan, sedangkan penyaringan relevansi domain dan panjang teks menghasilkan 15.155 ulasan.

Preprocessing mencakup normalisasi entitas HTML, penghapusan URL, mention, emoji, ekspresi tawa, serta karakter yang tidak diperlukan. Teks diubah menjadi huruf kecil, karakter berulang dinormalisasi, dan kata tidak baku, singkatan, slang, serta variasi bahasa daerah dinormalisasi menggunakan kamus normalisasi. Istilah yang berkaitan dengan GoFood dan variasi penulisan nominal juga ditangani dalam proses ini. Token residual yang belum tercakup dalam kamus diperiksa agar kata yang masih relevan tidak otomatis dihapus. Penelitian ini tidak menerapkan stemming maupun stopword removal, sedangkan kata negasi, seperti "tidak", dipertahankan karena dapat memengaruhi makna sentimen.

Penyaringan dilakukan berdasarkan aturan relevansi dan karakteristik teks. Meskipun membantu memusatkan analisis pada ulasan yang sesuai dengan domain penelitian, aturan tersebut berpotensi melewatkan ulasan relevan yang menggunakan kata di luar daftar kata kunci. Keterbatasan ini dipertimbangkan dalam interpretasi hasil penelitian.

2.4. Ekstraksi Aspek dan Pelabelan Sentimen

Ekstraksi aspek dilakukan menggunakan pendekatan *Rule-Based* berbasis kamus untuk mengidentifikasi empat aspek layanan GoFood, yaitu kualitas makanan, harga, pelayanan, dan aplikasi. Kamus aspek disusun berdasarkan kata kunci yang merepresentasikan setiap kategori, kemudian dicocokkan dengan teks ulasan yang telah melalui tahap *preprocessing*. Untuk menangani ulasan yang memuat lebih dari satu aspek, teks disegmentasikan menggunakan kata penghubung, seperti "tapi", "namun", "tetapi", "meskipun", "padahal", "dan", serta "sedangkan". Variasi bahasa informal, termasuk slang, singkatan, dan variasi ejaan, ditangani melalui proses normalisasi pada tahap persiapan data.

Penggunaan kamus aspek memungkinkan identifikasi aspek berdasarkan kata kunci yang telah ditentukan. Namun, pendekatan ini memiliki keterbatasan karena kata atau ungkapan yang tidak tercakup dalam kamus berpotensi tidak terdeteksi. Oleh karena itu, hasil ekstraksi aspek diperiksa melalui audit dan anotasi manual untuk mengevaluasi kesesuaian aspek yang teridentifikasi dengan konteks ulasan. Proses pemeriksaan ini digunakan untuk mengidentifikasi ketidaksesuaian hasil ekstraksi dan menilai kualitas aspek yang dihasilkan.

Pelabelan sentimen awal dilakukan menggunakan model *w11wo/indonesian-roberta-base-sentiment-classifier* untuk mengelompokkan ulasan ke dalam tiga kelas, yaitu positif, netral, dan negatif. Prediksi model digunakan sebagai label awal, bukan sebagai label final, karena sentimen global suatu ulasan belum tentu merepresentasikan sentimen terhadap aspek tertentu. Oleh sebab itu, hasil prediksi diperiksa melalui audit blind dan verifikasi manual dengan mempertimbangkan konteks aspek yang dibahas dalam ulasan.

Audit dilakukan terhadap 350 sampel dan menghasilkan akurasi sebesar 55,43% serta Macro F1-score sebesar 0,5319. Hasil tersebut menunjukkan bahwa prediksi awal RoBERTa masih memiliki keterbatasan sehingga tidak digunakan secara langsung sebagai label final. Label yang telah diverifikasi secara manual kemudian digunakan dalam pembentukan dataset ABSA. Dataset akhir terdiri atas 5.657 baris dari 4.786 *tweet* unik yang mencakup empat aspek dan tiga kelas sentimen. Dataset tersebut selanjutnya digunakan dalam proses pelatihan dan evaluasi model IndoBERT.

2.5. Transformasi Data dan Pelatihan Model

Dataset ABSA dibagi menjadi data pelatihan, validasi, dan pengujian dengan proporsi target 70:15:15. Pembagian dilakukan menggunakan *StratifiedGroupKFold* dengan 20 kelompok dan *random_state*=42. Pengelompokan didasarkan pada identitas *tweet* (*source_id*) agar pasangan aspek dari *tweet* yang sama tidak tersebar ke subset pelatihan, validasi, dan pengujian. Stratifikasi mempertimbangkan kombinasi aspek dan label sentimen untuk membantu menjaga distribusi kategori pada setiap subset.

Sebelum pelatihan, data ditransformasikan menggunakan tokenizer dari model *indobenchmark/indobert-base-p2*. Masukan model berbentuk pasangan teks yang terdiri atas nama aspek dan *text window*, yaitu potongan teks di sekitar kata kunci aspek yang digunakan dalam proses anotasi. Panjang masukan maksimum ditetapkan sebanyak 96 token. Format pasangan ini membantu model menentukan sentimen yang diminta berdasarkan aspek yang sedang dianalisis.

Model dilatih melalui proses *fine-tuning*, yaitu penyesuaian parameter model pralatih untuk tugas klasifikasi sentimen pada dataset penelitian. Pelatihan menggunakan 100 *epoch* tanpa *early stopping*, sesuai konfigurasi eksperimen yang digunakan. Parameter pelatihan dipilih berdasarkan percobaan dan kinerja pada data validasi. *Learning rate* sebesar 1×10^{-5} digunakan untuk mengatur besarnya pembaruan bobot model, sedangkan *weight decay* sebesar 0,08 dan *dropout* sebesar 0,25 digunakan sebagai mekanisme regularisasi untuk membantu mengurangi risiko *overfitting*. *Warmup ratio* sebesar 0,10 dan *linear learning rate scheduler* digunakan untuk mengatur perubahan laju pembelajaran selama pelatihan.

Pelatihan menggunakan *batch size* 32, *batch size* validasi 64, dan *gradient accumulation* sebanyak empat langkah. Dengan demikian, akumulasi gradien memungkinkan pembaruan bobot dilakukan setelah gradien dari beberapa langkah terakumulasi. *Optimizer AdamW* digunakan untuk memperbarui parameter model. *Embedding layer* dan enam *encoder layer* pertama dibekukan, sehingga pembaruan parameter difokuskan pada enam *encoder layer* bagian atas dan lapisan klasifikasi. Konfigurasi ini digunakan untuk menyesuaikan pelatihan model dengan kebutuhan eksperimen. Pemilihan checkpoint akhir didasarkan pada nilai *macro F1-score* terbaik pada data validasi, sedangkan data pengujian tidak digunakan untuk memilih *checkpoint*. Konfigurasi model yang digunakan pada penelitian ini ditunjukkan pada Tabel 1.

Tabel 1. Konfigurasi Model IndoBERT

Parameter	Nilai
Jumlah *Epoch*	100
Batch Size (*Training*)	32
Batch Size (*Validation*)	64
Gradient Accumulation	4
Learning Rate	1×10^{-5}

Weight Decay	0,08
Warmup Rasio	0,10
Learning Rate Scheduler	Linear
Hidden Dropout	0,25
Attention Dropout	0,25
Loss Mode	class_weight
Loss Function	*Weighted Cross-Entropy Loss*
Nilai Gamma (γ)	2,0
Freeze Layer	*Embedding Layer* + 6 *Encoder layer* Pertama
Optimizer	AdamW

Konfigurasi hiperparameter pada Tabel 1 dipilih dengan mempertimbangkan hasil percobaan *fine-tuning* dan kinerja model pada data validasi. Evaluasi pada data validasi digunakan sebagai dasar untuk menentukan konfigurasi yang digunakan dalam pelatihan akhir. Model yang telah dilatih selanjutnya dievaluasi menggunakan data pengujian untuk mengukur kemampuan klasifikasi sentimen terhadap data yang tidak digunakan selama pelatihan.

2.6. Evaluasi Model

Evaluasi dilakukan untuk mengukur kinerja ekstraksi aspek dan klasifikasi sentimen secara terpisah. Evaluasi ekstraksi aspek dilakukan dengan membandingkan aspek hasil *Rule-Based* terhadap anotasi manual. Pengukuran mencakup kesesuaian hasil identifikasi dan pemeriksaan pola kesalahan untuk mengetahui aspek yang sering salah dikenali. Karena kesepakatan anotasi dapat memengaruhi interpretasi hasil, hasil evaluasi ekstraksi juga dipertimbangkan bersama pemeriksaan kesepakatan antar-pelabel.

Evaluasi klasifikasi sentimen dilakukan pada data pengujian menggunakan *confusion matrix*, *accuracy*, *precision*, *recall*, dan *F1-score*. *Accuracy* mengukur proporsi prediksi benar dari seluruh data, *precision* mengukur ketepatan prediksi suatu kelas, *recall* mengukur kemampuan model mengenali data aktual pada kelas tersebut, sedangkan *F1-score* menggabungkan *precision* dan *recall*. Evaluasi dilaporkan secara keseluruhan dan berdasarkan empat aspek layanan. Nilai *macro average* digunakan untuk melihat rata-rata kinerja antarkelas dengan bobot sama, sedangkan *weighted average* mempertimbangkan jumlah data pada setiap kelas.

Selain mengevaluasi IndoBERT menggunakan aspek yang telah diverifikasi, pemeriksaan sistem secara menyeluruh dilakukan dengan mengalirkan hasil ekstraksi aspek *Rule-Based* ke klasifikasi sentimen IndoBERT. Evaluasi ini diperlukan untuk melihat kinerja alur ABSA secara lebih utuh, karena kesalahan ekstraksi aspek dapat memengaruhi hasil akhir klasifikasi sentimen. Seluruh evaluasi menggunakan label manual yang telah diverifikasi sebagai acuan.

3. Hasil dan Pembahasan

3.1. Hasil Ekstraksi Aspek

Ekstraksi aspek dilakukan menggunakan pendekatan *Rule-Based* berdasarkan kamus aspek yang telah disusun untuk mengidentifikasi pembahasan mengenai layanan GoFood. Penelitian ini menggunakan empat aspek, yaitu harga, pelayanan, aplikasi, dan kualitas makanan. Setiap aspek yang teridentifikasi dipasangkan dengan segmen teks yang relevan sehingga satu ulasan dapat menghasilkan lebih dari satu pasangan aspek–sentimen apabila membahas beberapa aspek sekaligus.

Berdasarkan hasil pemrosesan dan verifikasi manual, dataset akhir terdiri atas 5.657 pasangan aspek–sentimen yang berasal dari 4.786 *tweet* unik. Distribusi data berdasarkan aspek ditunjukkan pada Tabel 2.

Tabel 2. Distribusi Aspek pada Dataset Akhir

Aspek	Jumlah Data	Persentase

Harga	2.548	45,0%
Pelayanan	1.175	20,8%
Kualitas Makanan	989	17,5%
Aplikasi	945	16,7%
Total	5.657	100%

Berdasarkan Tabel 2, aspek harga memiliki jumlah data terbanyak, yaitu 2.548 data atau sekitar 45,04% dari keseluruhan dataset. Aspek pelayanan menempati urutan berikutnya dengan 1.175 data, diikuti kualitas makanan sebanyak 989 data dan aplikasi sebanyak 945 data. Distribusi ini menunjukkan bahwa pembahasan mengenai harga paling banyak ditemukan pada dataset ulasan yang digunakan dalam penelitian. Namun, frekuensi tersebut tidak secara langsung menunjukkan bahwa harga merupakan aspek dengan tingkat kepuasan atau ketidakpuasan tertinggi karena penentuan sentimennya dianalisis secara terpisah.

Hasil ekstraksi awal selanjutnya diperiksa melalui audit, inspeksi, dan verifikasi manual. Pemeriksaan dilakukan dengan mempertimbangkan konteks ulasan menggunakan kolom *text_window* sebagai konteks utama dan *text_segment* sebagai konteks tambahan apabila diperlukan. Tahapan ini diperlukan karena pencocokan kata kunci berpotensi menghasilkan aspek yang kurang sesuai dengan konteks atau tidak mengenali variasi ungkapan yang digunakan pengguna.

3.2. Evaluasi Ekstraksi Aspek Rule-Based

Evaluasi ekstraksi aspek dilakukan dengan membandingkan aspek yang dihasilkan oleh metode *Rule-Based* dengan label aspek manual pada 5.690 data audit. Berdasarkan hasil, pendekatan *Rule-Based* memperoleh akurasi sebesar 94,66% dan *Macro F1-score* sebesar 0,9324. Rincian hasil evaluasi ditunjukkan pada Tabel 3.

Tabel 3. Hasil Evaluasi Ekstraksi Aspek Rule-Based

Aspek	***Precision***	***Recall***	***F1-Score***	***Support***
Aplikasi	0,9973	0,7470	0,8542	1.004
Harga	0,9857	0,9885	0,9871	2.518
Kualitas Makanan	0,9828	0,9775	0,9801	933
Pelayanan	0,8316	1,0000	0,9081	1.235
Macro Average	0,9494	0,9282	0,9324	5.690
Accuracy			0,9466	5.690

Berdasarkan Tabel 3, aspek harga memperoleh *F1-score* tertinggi sebesar 0,9871, diikuti kualitas makanan sebesar 0,9801. Hasil tersebut menunjukkan kesesuaian yang tinggi antara hasil identifikasi kedua aspek tersebut dan label manual pada data audit. Aspek pelayanan memperoleh *recall* sebesar 1,0000, yang berarti seluruh data berlabel pelayanan pada sampel audit berhasil dikenali sebagai pelayanan oleh metode *Rule-Based*. Namun, *precision*-nya sebesar 0,8316 menunjukkan bahwa sebagian data yang diprediksi sebagai pelayanan memiliki label manual aspek lain.

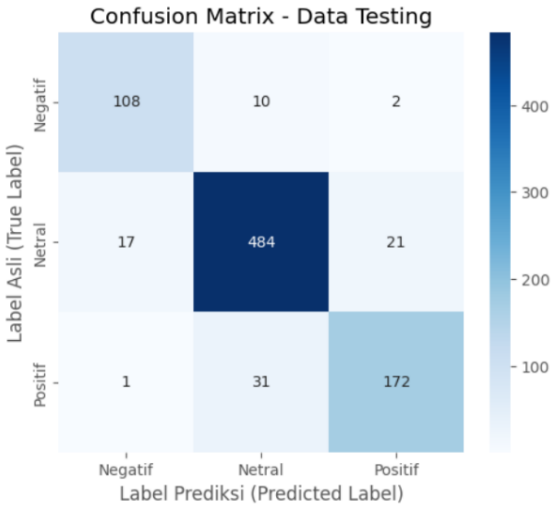
Sementara itu, aspek aplikasi memperoleh nilai *recall* sebesar 0,7470, lebih rendah dibandingkan aspek lainnya. Hasil ini menunjukkan bahwa metode *Rule-Based* belum mampu mengidentifikasi seluruh data yang termasuk dalam aspek aplikasi secara optimal. Salah satu bentuk kesalahan yang ditemukan adalah ulasan yang seharusnya termasuk dalam aspek aplikasi justru teridentifikasi sebagai aspek pelayanan. Hal ini mengindikasikan bahwa penggunaan kata kunci tertentu, seperti penyebutan GoFood, belum selalu dapat membedakan pembahasan mengenai aplikasi dari pembahasan mengenai pelayanan.

Meskipun evaluasi menunjukkan akurasi sebesar 94,66%, hasil tersebut perlu diinterpretasikan dengan mempertimbangkan proses anotasi yang digunakan. Pada evaluasi terhadap 600 data yang melibatkan pelabel kedua, tingkat kesesuaian hasil *Rule-Based* dengan label pelabel kedua mencapai 41,67%, atau 44,64% setelah data yang dinyatakan tidak relevan dikeluarkan. Sebaliknya, kesesuaian antara hasil *Rule-Based* dan label peneliti mencapai 94,17% pada sampel yang sama. Perbedaan ini

menunjukkan adanya kemungkinan pengaruh perbedaan interpretasi dan penerapan pedoman anotasi terhadap hasil evaluasi. Oleh karena itu, nilai akurasi perlu dipahami dalam konteks proses pelabelan yang dilakukan dan tidak dapat dijadikan satu-satunya dasar untuk menyimpulkan ketepatan ekstraksi aspek secara independen.

3.3. Hasil Klasifikasi Sentimen Menggunakan IndoBERT

Tahap pengujian dilakukan setelah model IndoBERT menjalani proses *fine-tuning*. Dataset pengujian digunakan untuk menilai kemampuan model dalam mengklasifikasikan sentimen pada setiap aspek yang telah diekstraksi. Penilaian kinerja model didasarkan pada *confusion matrix* serta empat metrik evaluasi, yaitu *accuracy*, *precision*, *recall*, dan *F1-score*. Hasil pengujian dalam bentuk *confusion matrix* disajikan dalam Gambar 2, sedangkan hasil *classification report* disajikan pada Tabel 4.



Gambar 2. Confusion Matrix Hasil Klasifikasi Sentimen IndoBERT

Berdasarkan Gambar 2, model IndoBERT berhasil mengklasifikasikan 108 dari 120 data negatif, 484 dari 522 data netral, dan 172 dari 204 data positif secara benar. Secara keseluruhan, sebanyak 764 dari 846 data pengujian berhasil diklasifikasikan dengan tepat, sedangkan 82 data lainnya mengalami kesalahan klasifikasi. Kesalahan terbanyak terjadi pada data bersentimen positif yang diprediksi sebagai netral, yaitu sebanyak 31 data. Selain itu, terdapat 21 data netral yang diprediksi sebagai positif, 17 data netral yang diprediksi sebagai negatif, dan 10 data negatif yang diprediksi sebagai netral. Hasil ini menunjukkan bahwa model masih mengalami kesulitan dalam membedakan sebagian ulasan positif dan netral. Meskipun demikian, jumlah prediksi benar pada setiap kelas menunjukkan bahwa model mampu mengklasifikasikan ketiga kelas sentimen dengan cukup baik pada data pengujian.

Tabel 4. Hasil Evaluasi Klasifikasi Sentimen IndoBERT

Label	***Precision***	***Recall***	***F1-score***	***Support***
Negatif	0,86	0,90	0,88	120
Netral	0,92	0,93	0,92	522
Positif	0,88	0,84	0,86	204
Macro Average	0,89	0,89	0,89	846
Weighted Average	0,90	0,90	0,90	846

Pada Tabel 4, kelas netral memperoleh *F1-score* tertinggi sebesar 0,9245, dengan *precision* 0,9219 dan *recall* 0,9272. Hasil ini menunjukkan bahwa model mampu mengklasifikasikan sentimen netral dengan baik pada data pengujian. Kelas negatif memperoleh *F1-score* sebesar 0,8780, sedangkan kelas positif memperoleh *F1-score* sebesar 0,8622. Dengan demikian, kinerja model berbeda pada setiap kelas sentimen dan tidak hanya ditentukan oleh nilai akurasi keseluruhan.

Jumlah data pengujian juga tidak seimbang antarkelas. Kelas netral memiliki 522 data, sedangkan kelas negatif dan positif masing-masing memiliki 120 dan 204 data. Oleh karena itu, *Macro F1-score* sebesar 0,8883 perlu diperhatikan bersama metrik per kelas karena memberikan bobot yang sama kepada setiap kelas, sedangkan *Weighted F1-score* mempertimbangkan jumlah data pada masing-masing kelas.

3.4. Analisis Klasifikasi Sentimen Berdasarkan Aspek

Pengujian selanjutnya dilakukan untuk menilai performa model IndoBERT dalam mengklasifikasikan sentimen pada masing-masing aspek layanan. Evaluasi ini berfokus pada kemampuan klasifikasi sentimen untuk setiap aspek yang dianalisis. Hasil evaluasi disajikan pada Tabel 5.

Tabel 5. Hasil Klasifikasi Sentimen IndoBERT Berdasarkan Aspek

Aspek	Jumlah Data Uji	*Accuracy*	*Macro F1-Score*
Kualitas Makanan	149	92,62%	0,8975
Harga	375	90,40%	0,8841
Pelayanan	190	90,00%	0,8352
Aplikasi	132	87,88%	0,7338
Keseluruhan	846	90,31%	0,8883

Berdasarkan Tabel 5, aspek kualitas makanan memperoleh akurasi tertinggi sebesar 92,62% dan *Macro F1-score* sebesar 0,8975. Hasil ini menunjukkan bahwa model memiliki kinerja relatif baik dalam mengklasifikasikan sentimen pada aspek kualitas makanan di data pengujian. Aspek harga memperoleh akurasi sebesar 90,40% dan *Macro F1-score* 0,8841, sedangkan aspek pelayanan memperoleh akurasi 90,00% dan *Macro F1-score* 0,8352.

Aspek aplikasi memperoleh akurasi sebesar 87,88% dan *Macro F1-score* sebesar 0,7338, yaitu *Macro F1-score* terendah di antara empat aspek. Hasil ini berkaitan dengan perbedaan kinerja antarkelas pada aspek tersebut. Kelas negatif hanya memiliki tujuh data pengujian, sementara kelas netral memiliki 103 data dan kelas positif 22 data. Dengan jumlah sampel negatif yang sangat kecil, metrik per kelas terutama untuk kelas negatif perlu ditafsirkan secara hati-hati dan belum tentu stabil apabila diuji pada sampel lain.

Pada aspek kualitas makanan, terdapat 90 data positif, 45 data netral, dan 14 data negatif. *F1-score* kelas positif mencapai 0,9617, sedangkan kelas negatif dan netral masing-masing mencapai 0,8571 dan 0,8736. Hasil tersebut berkontribusi terhadap *Macro F1-score* aspek kualitas makanan sebesar 0,8975.

Dengan demikian, hasil evaluasi per aspek memperlihatkan bahwa akurasi yang tinggi tidak selalu diikuti oleh nilai *Macro F1-score* yang sama tingginya. Oleh sebab itu, pembahasan kinerja model perlu mempertimbangkan distribusi kelas dan nilai evaluasi masing-masing kelas, bukan hanya membandingkan nilai akurasi.

3.5. Analisis Sentimen dan Implikasi terhadap Layanan GoFood

Aspek pelayanan memiliki proporsi sentimen negatif tertinggi, yaitu 18,9%, dengan sentimen netral sebesar 71,1%. Dibandingkan dengan aspek lain yang dianalisis, pelayanan memiliki proporsi ulasan negatif yang relatif lebih besar. Oleh karena itu, aspek pelayanan dapat menjadi salah satu prioritas dalam evaluasi layanan GoFood. Penyedia layanan dan mitra restoran dapat menelaah ulasan negatif secara lebih terperinci untuk mengidentifikasi jenis keluhan yang muncul. Hasil penelaahan tersebut dapat digunakan sebagai bahan evaluasi terhadap proses pelayanan, komunikasi dengan pelanggan, dan penanganan kendala pesanan. Dengan demikian, rekomendasi perbaikan dapat didasarkan pada permasalahan yang tercermin dalam ulasan pengguna.

Pada aspek harga, sentimen netral mendominasi dengan proporsi sebesar 63,7%, sedangkan pada aspek aplikasi proporsinya mencapai 78,0%. Dominasi sentimen netral menunjukkan bahwa sebagian besar ulasan pada kedua aspek tersebut diklasifikasikan sebagai netral. Namun, kondisi ini tidak dapat langsung diartikan sebagai kepuasan maupun ketidakpuasan pengguna. Oleh karena itu, isi ulasan netral perlu ditelaah lebih lanjut untuk memahami konteks pengalaman pengguna, termasuk pembahasan mengenai harga dan penggunaan aplikasi. Penelaahan tersebut dapat membantu mengidentifikasi isu yang perlu dievaluasi tanpa mengasumsikan adanya permasalahan yang belum didukung oleh data.

Sementara itu, aspek kualitas makanan menunjukkan proporsi sentimen positif tertinggi, yaitu 60,4%. Hal ini menunjukkan bahwa ulasan mengenai kualitas makanan dalam data yang dianalisis lebih banyak diklasifikasikan sebagai positif dibandingkan sebagai negatif atau netral. Temuan tersebut dapat menjadi informasi pendukung bagi GoFood dan mitra restoran untuk mempertahankan kualitas produk yang memperoleh penilaian positif dari pengguna. Jika dibandingkan dengan aspek pelayanan, hasil ini juga memperlihatkan perbedaan kecenderungan sentimen antara kualitas makanan dan pelayanan dalam data pengujian.

Secara keseluruhan, hasil analisis sentimen berbasis aspek dapat menjadi informasi pendukung dalam menentukan aspek layanan yang perlu dipertahankan atau ditinjau lebih lanjut. Aspek pelayanan dapat diprioritaskan untuk penelaahan ulasan negatif, sedangkan ulasan mengenai harga dan aplikasi dapat dianalisis lebih lanjut untuk memahami konteks pengalaman pengguna. Meskipun demikian, hasil penelitian ini merepresentasikan kecenderungan sentimen dalam ulasan yang dianalisis dari platform X dan tidak dapat digeneralisasikan secara langsung kepada seluruh pengguna GoFood. Penentuan prioritas perbaikan juga perlu mempertimbangkan isi ulasan secara lebih terperinci, bukan hanya proporsi sentimen.

3.6. Analisis Kesalahan Klasifikasi Model

Analisis kesalahan klasifikasi dilakukan berdasarkan *confusion matrix* untuk mengidentifikasi pola prediksi yang belum tepat pada model IndoBERT. Dari 846 data uji, sebanyak 764 data berhasil diklasifikasikan dengan benar, sedangkan 82 data mengalami kesalahan klasifikasi. Hasil tersebut menunjukkan bahwa meskipun model memperoleh *accuracy* sebesar 90,31%, masih terdapat ulasan yang belum dapat diklasifikasikan secara tepat. Oleh karena itu, analisis terhadap kesalahan prediksi diperlukan untuk memahami karakteristik klasifikasi yang masih menjadi tantangan bagi model.

Pada kelas negatif, sebanyak 108 dari 120 data berhasil diprediksi dengan benar, sedangkan 10 data diprediksi sebagai netral dan 2 data sebagai positif. Pada kelas netral, sebanyak 484 dari 522 data diklasifikasikan dengan benar, sementara 17 data diprediksi sebagai negatif dan 21 data sebagai positif. Adapun pada kelas positif, sebanyak 172 dari 204 data berhasil diklasifikasikan dengan benar, sedangkan 31 data diprediksi sebagai netral dan 1 data sebagai negatif. Pola tersebut menunjukkan bahwa kesalahan prediksi tidak tersebar secara merata. Kesalahan paling banyak terjadi pada data positif yang diprediksi sebagai netral, diikuti oleh data netral yang diprediksi sebagai positif dan data negatif yang diprediksi sebagai netral.

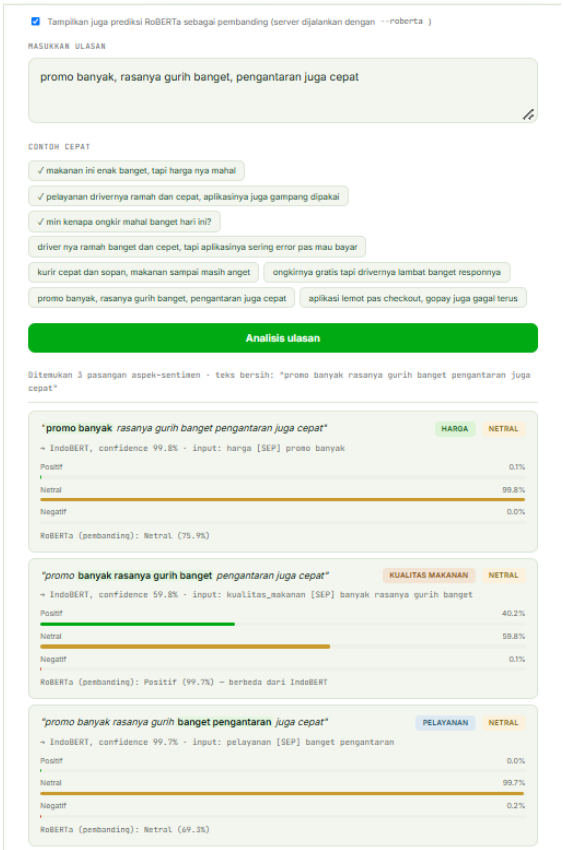
Kesalahan antara kelas negatif dan netral menjadi salah satu hal yang perlu diperhatikan dalam klasifikasi sentimen ulasan GoFood. Ulasan pada platform X dapat memuat informasi mengenai pengalaman pengguna atau kendala layanan tanpa selalu menggunakan ungkapan negatif secara eksplisit. Sebaliknya, penyampaian informasi mengenai suatu pengalaman tidak selalu menunjukkan ketidakpuasan. Perbedaan konteks tersebut berpotensi menyulitkan penentuan sentimen, terutama ketika suatu ulasan tidak menyampaikan evaluasi secara langsung. Namun, pola pada *confusion matrix* saja belum cukup untuk memastikan bahwa karakteristik tersebut merupakan penyebab kesalahan prediksi. Verifikasi lebih lanjut melalui pemeriksaan teks ulasan yang salah diklasifikasikan diperlukan untuk mengidentifikasi penyebabnya secara lebih tepat.

Kesalahan lain yang menonjol terjadi pada pembedaan sentimen positif dan netral. Sebanyak 31 data positif diprediksi sebagai netral, sedangkan 21 data netral diprediksi sebagai positif. Pola ini menunjukkan bahwa model masih mengalami kesulitan dalam membedakan sebagian ulasan yang mengandung penilaian positif dari ulasan yang cenderung bersifat informatif. Karakteristik bahasa informal, penggunaan ungkapan evaluatif yang tidak eksplisit, serta konteks kalimat dapat menjadi faktor yang perlu diperiksa dalam analisis lanjutan. Akan tetapi, faktor-faktor tersebut masih berupa kemungkinan dan belum dapat dinyatakan sebagai penyebab pasti tanpa pemeriksaan terhadap contoh ulasan yang mengalami kesalahan klasifikasi.

Temuan ini memperlihatkan bahwa penilaian kinerja model tidak cukup hanya mengandalkan nilai *accuracy*, tetapi juga perlu mempertimbangkan distribusi dan pola kesalahan pada setiap kelas sentimen. Meskipun model menunjukkan kinerja yang baik secara keseluruhan, kesalahan antara kelas negatif dan netral serta antara kelas positif dan netral menunjukkan adanya ruang untuk perbaikan. Pemeriksaan kualitatif terhadap ulasan yang salah diklasifikasikan dapat membantu mengidentifikasi karakteristik teks yang berkaitan dengan kesalahan prediksi. Selanjutnya, penelitian dapat mengevaluasi kebutuhan penambahan data beranotasi yang relevan dan peningkatan konsistensi anotasi untuk mendukung kemampuan model dalam membedakan sentimen yang memiliki karakteristik serupa.

3.7. Implementasi Aplikasi Web dan Keterbatasan Penelitian

Model *Aspect-Based Sentiment Analysis* (ABSA) yang dikembangkan dalam penelitian ini diimplementasikan ke dalam aplikasi berbasis web sebagai media demonstrasi hasil model. Aplikasi memungkinkan pengguna memasukkan ulasan layanan GoFood untuk dianalisis melalui proses ekstraksi aspek menggunakan pendekatan *Rule-Based* dan klasifikasi sentimen menggunakan IndoBERT. Selain itu, aplikasi menyediakan prediksi RoBERTa sebagai pembanding terhadap hasil klasifikasi IndoBERT. Tampilan halaman demo disajikan pada Gambar 3.



Gambar 3. Tampilan Halaman Demo Aplikasi Web

Berdasarkan Gambar 3, pengguna dapat memasukkan ulasan GoFood untuk dianalisis. Sistem menampilkan pasangan aspek dan segmen teks hasil ekstraksi, kemudian menghasilkan klasifikasi sentimen beserta nilai *confidence* dan distribusi probabilitas. Prediksi RoBERTa juga ditampilkan untuk memperlihatkan perbedaan keluaran kedua model. Implementasi ini berpotensi membantu penelaahan ulasan berdasarkan aspek layanan, tetapi masih terbatas sebagai demonstrasi karena belum dilakukan pengujian pengguna maupun evaluasi manfaatnya dalam penggunaan nyata.

Penelitian ini memiliki beberapa keterbatasan. Data yang berasal dari platform X belum tentu mewakili seluruh pengguna GoFood sehingga hasilnya tidak dapat digeneralisasikan secara langsung. Selain itu, ekstraksi aspek bergantung pada kamus dan aturan *Rule-Based*, yang dapat membatasi pengenalan variasi kosakata dan bahasa informal. Kinerja klasifikasi juga perlu dipahami berdasarkan distribusi dan karakteristik data pengujian. Penelitian selanjutnya dapat memperluas variasi data, menyempurnakan aturan ekstraksi aspek, menganalisis kesalahan klasifikasi, serta melakukan pengujian pengguna dan evaluasi komparatif IndoBERT–RoBERTa menggunakan data uji dan metrik yang sama.

4. Kesimpulan

Penelitian ini menerapkan *Aspect-Based Sentiment Analysis* (ABSA) menggunakan pendekatan hibrida, yaitu *Rule-Based* untuk ekstraksi aspek dan IndoBERT untuk klasifikasi sentimen pada ulasan GoFood di platform X. Dari 28.763 data awal, diperoleh 5.657 pasangan aspek–sentimen yang telah melalui proses audit dan verifikasi manual. Penelitian ini mengidentifikasi empat aspek layanan, yaitu harga, kualitas makanan, pelayanan, dan aplikasi. Pada 846 data pengujian, model IndoBERT memperoleh akurasi sebesar 90,31% dan Macro F1-score sebesar 0,8883. Hasil ini menunjukkan bahwa model memiliki kinerja yang baik dalam mengklasifikasikan sentimen pada dataset penelitian.

Secara praktis, hasil ABSA dapat menjadi informasi pendukung untuk memahami kecenderungan sentimen pengguna pada setiap aspek layanan dan membantu mengidentifikasi aspek yang perlu ditinjau lebih lanjut. Namun, penelitian ini belum mengevaluasi secara langsung penerapan hasil analisis dalam pengambilan keputusan maupun dampaknya terhadap peningkatan kualitas layanan GoFood.

Penelitian ini memiliki keterbatasan berupa ketergantungan ekstraksi aspek pada kamus kata kunci, yang berpotensi membatasi identifikasi variasi kosakata, bahasa tidak baku, dan konteks ulasan. Selain itu, data hanya berasal dari platform X dengan GoFood sebagai studi kasus dan mencakup empat aspek serta tiga kategori sentimen. Penelitian selanjutnya dapat mengembangkan metode ekstraksi aspek yang lebih adaptif, memperluas sumber dan cakupan data, serta membandingkan IndoBERT dengan model *Transformer* lain, seperti IndoRoBERTa dan XLM-RoBERTa. Pengujian terhadap pengguna dan evaluasi penerapan hasil analisis juga diperlukan untuk mengetahui kegunaan praktis sistem dalam mendukung evaluasi layanan.

Daftar Pustaka

[1] N. S. Shaeeali, A. Mohamed, and S. Mutalib, "Customer reviews analytics on food delivery services in social media: A review," *IAES Int. J. Artif. Intell.*, vol. 9, no. 4, pp. 691–699, 2020, doi: 10.11591/ijai.v9.i4.pp691-699.

[2] A. Fathrizqy, P. Mahardika, and A. Muid, "The Sentiment Analysis of Online Customer Review on Food and Beverages Delivery Services in the GOJEK Application Using K-Nearest Neighbors," pp. 854–862, 2023, doi: 10.46254/ap03.20220167.

[3] P. J. J. S. Gautama, A. Wibowo, and J. J. Siang, "Analisis Sentimen Pengguna Terhadap Akun X/Twitter Resmi 'DANA' Dengan Algoritma Indobert," *INTECOMS J. Inf. Technol. Comput. Sci.*, vol. 8, no. 2, pp. 549–559, 2025, doi: 10.31539/intecoms.v8i2.14477.

[4] F. Aftab *et al.*, "A Comprehensive Survey on Sentiment Analysis Techniques," *Int. J. Technol.*, vol. 14, no. 6, pp. 1288–1298, 2023, doi: 10.14716/ijtech.v14i6.6632.

[5] G. Brauwers and F. Frasincar, "A Survey on Aspect-Based Sentiment Classification," *ACM Comput. Surv.*, vol. 55, no. 4, pp. 1–35, 2023, doi: 10.1145/3503044.

[6] R. Chauhan, A. Gusain, P. Kumar, C. Bhatt, and I. Uniyal, "Fine Grained Sentiment Analysis using Machine Learning and Deep Learning," *Int. Conf. Sustain. Emerg. Innov. Eng. Technol. ICSEIET 2023*, vol. 11, no. 5s, pp. 423–427, 2023, doi: 10.1109/ICSEIET58677.2023.10303481.

[7] Anushree Goud and Bindu Garg, "Advancements in Aspect-Based Sentiment Analysis: Leveraging Deep Learning and Machine Learning Techniques," *J. Inf. Syst. Eng. Manag.*, vol. 10, no. 21s, pp. 451–473, 2025, doi: 10.52783/jisem.v10i21s.3371.

[8] J. Siswanto, S. Suakanto, M. Andriani, M. Hardiyanti, and T. F. Kusumasari, "Interview Bot Development with Natural Language Processing and Machine Learning," *Int. J. Technol.*, vol. 13, no. 2, pp. 274–285, 2022, doi: 10.14716/ijtech.v13i2.5018.

[9] C. Yusni Adhelina, M. Halmi Dar, and M. Nirmala Sari Hasibuan, "Implementation of Deep Learning Models in Conducting Aspect-Based Sentiment Analysis," *Int. J. Sci. Technol. Manag.*, vol. 6, no. 4, pp. 949–957, 2025, doi: 10.46729/ijstm.v6i4.1251.

[10] A. Wafda, D. H. Fudholi, and J. Nugraha, "Aspect-Based Sentiment Analysis on Twitter Tweets About the Merdeka Curriculum Using Indobert," *JITK (Jurnal Ilmu Pengetah. dan Teknol. Komputer)*, vol. 10, no. 3, pp. 586–599, 2025, doi: 10.33480/jitk.v10i3.5692.

[11] M. Rayie, R. Putera, and H. Irawan, "Analisis Sentimen Ulasan Produk Toko Online Rava.Id Untuk Peningkatan Produk Menggunakan Algoritma IndoBERT-RoBERTa," Universitas Telkom, 2025. [Online]. Available: https://repository.telkomuniversity.ac.id/pustaka/239729/

[12] A. Jazuli, Widowati, and R. Kusumaningrum, "Optimizing Aspect-Based Sentiment Analysis Using BERT for Comprehensive Analysis of Indonesian Student Feedback," *Appl. Sci.*, vol. 15, no. 1, pp. 1–28, 2025, doi: 10.3390/app15010172.

[13] V. D. Setiawan, D. U. Iswavigra, and E. Anggiratih, "Implementation of IndoBERT for Sentiment Analysis of the Constitutional Court's Decision Regarding the Minimum Age of Vice Presidential Candidates," *Sci. J. Informatics*, vol. 12, no. 3, pp. 397–406, 2025, doi: 10.15294/sji.v12i3.26320.

[14] F. Koto, A. Rahimi, J. H. Lau, and T. Baldwin, "IndoLEM and IndoBERT: A Benchmark Dataset and Pre-trained Language Model for Indonesian NLP," *COLING 2020 - 28th Int. Conf. Comput. Linguist. Proc. Conf.*, pp. 757–770, 2020, doi: 10.18653/v1/2020.coling-main.66.

[15] Anugerah Simanjuntak *et al.*, "Research and Analysis of IndoBERT Hyperparameter Tuning in Fake News Detection," *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 13, no. 1, pp. 60–67, 2024, doi: 10.22146/jnteti.v13i1.8532.

[16] R. Talita Trista and E. Tri Asmoro, "Analisis Sentimen dalam Data Big Data," *J. Multidiscip. Inq. Sci. Technol. Educ. Res.*, vol. 2, no. 1, pp. 700–706, 2025.

[17] W. Ahmad, H. U. Khan, F. K. Alarfaj, and M. Alreshoodi, "Aspect-Based Sentiment Analysis: A Comprehensive Review and Open Research Challenges," *IEEE Access*, vol. 13, pp. 65138–65182, 2025, doi: 10.1109/ACCESS.2025.3555744.

[18] N. F. Adhim and N. Cahyono, "Optimization of IndoBERT for Sentiment Analysis of FOMO on Social Media Through Fine-tuning and Hybrid Labeling," *J. Appl. Informatics Comput.*, vol. 9, no. 6, pp. 3786–3797, 2025.

[19] A. Fatihaturrahmah and K. Ditha Tania, "A Hybrid Approach of Aspect-Based Sentiment Analysis and Knowledge Extraction for Evaluating Security Perceptions in Digital Payment Applications," *Sci. J. Informatics*, vol. 12, no. 4, pp. 635–650, 2025, doi: 10.15294/sji.v12i4.31557.

[20] C. H. Lin and U. Nuha, "Sentiment analysis of Indonesian datasets based on a hybrid deep-learning strategy," *J. Big Data*, vol. 10, no. 1, 2023, doi: 10.1186/s40537-023-00782-9.