

Implementasi Model IndoBERT Untuk Analisis Sentimen Berbasis Aspek Dan Visualisasinya Pada Dashboard Interaktif Ulasan Restoran

Valois Vicenti Sirait¹, Parasian D. P Silitonga^{2*}

^{1,2}Teknik Informatika, Fakultas Ilmu Komputer, Universitas Katolik Santo Thomas, Medan, Indonesia

Email: ¹valoissirait0@email.com, ²parasianirene@email.com

INFORMASI ARTIKEL

Histori artikel:

Kata Kunci:

Aspect-Based Sentiment
Analysis, Dasbor Interaktif,
IndoBERT, Sentence Pair
Classification, Ulasan
Restoran....

ABSTRAK

Abstract- The increase in the number of customer reviews on restaurant digital platforms generates unstructured data that is difficult to analyze manually. Additionally, a single review often contains multiple aspects with different sentiments, necessitating an approach capable of identifying sentiment for each aspect separately. This research aims to implement the IndoBERT architecture to perform Aspect-Based Sentiment Analysis (ABSA) and visualize the results through an interactive analytics dashboard. The model was developed using the Sentence Pair Classification (SPC) approach on 1,080 customer reviews from BPK Tesalonika, paired with five aspect categories, resulting in 5,400 data pairs. The experimental results show that the learning rate configuration of 5×10^{-5} yields the best performance. Class distribution imbalance was addressed using a Cost-Sensitive Learning approach with a Weighted Cross-Entropy function. The test results show that the model is capable of detecting aspects with an accuracy of 95.09% and an F1-Score of 92.08%, while sentiment classification achieves an accuracy of 90.00% and a Macro F1-Score of 85.56%. The model was then integrated into a Flask-based application, allowing the analysis results to be presented thru an interactive dashboard. The developed system is able to automatically present summaries of aspect and sentiment distribution thru the interactive dashboard, serving as a supporting medium for service evaluation and data-driven decision-making.

Abstrak- Peningkatan jumlah ulasan pelanggan pada platform digital restoran menghasilkan data tidak terstruktur yang sulit untuk dianalisis secara manual. Selain itu, satu ulasan sering kali memuat beberapa aspek dengan sentimen yang berbeda sehingga diperlukan pendekatan yang mampu mengidentifikasi sentimen pada setiap aspek secara terpisah. Penelitian ini bertujuan mengimplementasikan arsitektur IndoBERT untuk melakukan Aspect-Based Sentiment Analysis (ABSA) serta memvisualisasikan hasilnya melalui dashboard analitik interaktif. Model dikembangkan menggunakan pendekatan Sentence Pair Classification (SPC) terhadap 1.080 ulasan pelanggan BPK Tesalonika yang dipasangkan dengan lima kategori aspek sehingga menghasilkan 5.400 pasangan data. Hasil eksperimen menunjukkan bahwa konfigurasi learning rate 5×10^{-5} memberikan performa terbaik. Ketidakseimbangan distribusi kelas ditangani menggunakan pendekatan Cost-Sensitive Learning dengan fungsi Weighted Cross-Entropy. Hasil pengujian menunjukkan bahwa model mampu mendeteksi aspek dengan akurasi 95,09% dan F1-Score 92,08%, sedangkan klasifikasi sentimen memperoleh akurasi 90,00% dan Macro F1-Score 85,56%. Model selanjutnya diintegrasikan ke dalam aplikasi berbasis Flask sehingga hasil analisis dapat disajikan melalui dashboard interaktif, sistem yang dikembangkan mampu menyajikan ringkasan distribusi aspek dan sentimen pelanggan secara otomatis melalui dashboard

interaktif sebagai media pendukung evaluasi layanan dan pengambilan keputusan berbasis data.

*Copyright © 2026 LPPM - STMIK IKMI Cirebon
This is an open access article under the CC-BY license*

Penulis Korespondensi:

Parasian D. P Silitonga

Teknik Informatika, Fakultas Ilmu Komputer, Universitas Katolik Santo Thomas, Medan,
Indonesia

Email: parasianirene@email.com

1. Pendahuluan

Perkembangan platform digital telah mendorong pelanggan untuk secara aktif membagikan pengalaman mereka melalui ulasan daring pada berbagai layanan, termasuk Google Maps. Ulasan pelanggan kini telah menjadi sumber informasi utama dalam proses pengambilan keputusan konsumen sekaligus memberikan informasi penting bagi pemilik usaha, karena memuat opini mengenai berbagai aspek terhadap usahanya dan pengalaman pelanggannya secara keseluruhan. Informasi tersebut dapat dimanfaatkan untuk melakukan evaluasi peningkatan kualitas layanan dan mempertahankan kepuasan pelanggan (Aureliano-Silva et al., 2021; S. Li et al., 2022). Namun, jumlah ulasan yang terus meningkat menyebabkan proses analisis manual, seperti membaca ulasan satu persatu menjadi kurang efisien, memerlukan waktu yang lama, serta berpotensi menghasilkan interpretasi yang tidak konsisten. Oleh karena itu, diperlukan pendekatan berbasis *Natural Language Processing* (NLP) untuk mengotomatisasi analisis dan meningkatkan konsistensi hasil (Shin Bee et al., 2022).

Analisis sentimen merupakan salah satu pendekatan yang banyak digunakan untuk mengekstraksi opini dari data teks seperti ulasan secara otomatis (Liu, 2012). Meskipun demikian, analisis sentimen konvensional, seperti analisis sentimen tingkat dokumen, umumnya hanya menghasilkan satu label sentimen untuk keseluruhan ulasan sehingga belum mampu menggambarkan opini pelanggan terhadap aspek tertentu. Pada domain restoran, satu ulasan dapat memuat berbagai opini terhadap aspek yang berbeda. Kondisi ini menyebabkan representasi sentimen tunggal menjadi kurang memadai untuk menggambarkan persepsi pelanggan secara menyeluruh (Hua et al., 2024). Untuk mengatasi keterbatasan tersebut, dikembangkan *Aspect-Based Sentiment Analysis* (ABSA) untuk

Untuk bahasa Indonesia, Wilie et al. (2020) mengembangkan IndoBERT yang dilatih menggunakan korpus bahasa Indonesia berskala besar dan terbukti efektif dalam berbagai tugas pemrosesan bahasa Indonesia. Pada tugas ABSA, Sun et al. (2019) memperkenalkan pendekatan *Sentence Pair Classification* (SPC) yang memformulasikan analisis sentimen berbasis aspek sebagai tugas klasifikasi pasangan kalimat antara ulasan dan aspek target, sehingga model dapat mempelajari hubungan antara aspek dan konteks ulasan secara lebih efektif.

Penelitian sebelumnya menunjukkan bahwa kombinasi IndoBERT dan SPC mampu memberikan performa yang baik pada tugas

mengidentifikasi aspek yang dibahas sekaligus menentukan polaritas sentimen pada setiap aspek secara terpisah (Liu, 2012; Pontiki et al., 2014).

Pentingnya ABSA tidak hanya terbatas pada aspek teknis pengolahan bahasa, tetapi juga memiliki implikasi strategis bagi keberlangsungan suatu usaha. Penelitian yang dilakukan oleh H. Li et al. (2023) menunjukkan bahwa analisis sentimen berbasis aspek mampu memberikan informasi yang lebih akurat mengenai faktor-faktor yang memengaruhi kepuasan pelanggan dibandingkan analisis sentimen umum. Informasi yang lebih rinci mengenai aspek-aspek yang mendapatkan pujian maupun keluhan dari pelanggan memungkinkan pemilik usaha untuk memahami sumber kepuasan dan ketidakpuasan pelanggan secara lebih tepat. Lohith et al. (2025) menegaskan bahwa kemampuan mengidentifikasi sentimen pada setiap aspek sangat penting untuk membantu pihak manajemen menemukan akar permasalahan yang sebenarnya dan menyusun strategi untuk evaluasi perbaikan yang lebih efektif. Oleh karena itu, keberhasilan penerapan ABSA sangat bergantung pada kemampuan model dalam memahami konteks bahasa dan hubungan antara aspek dengan opini yang disampaikan pelanggan.

Seiring perkembangan teknologi *Natural Language Processing* (NLP), model berbasis Transformer menunjukkan performa yang unggul pada berbagai tugas klasifikasi teks. Arsitektur Transformer mampu memahami hubungan kontekstual antar kata melalui mekanisme *self-attention* (Vaswani et al., 2017). Salah satu implementasi Transformer yang banyak digunakan adalah *Bidirectional Encoder* makanan secara manual. Penelitian yang telah *Representations from Transformers* (BERT), yang mampu menghasilkan representasi bahasa secara dua arah sehingga lebih efektif dalam memahami konteks kalimat (Devlin et al., 2019).

ABSA berbahasa Indonesia. Yulianti et al. (2024) melaporkan bahwa pendekatan tersebut memperoleh F1-Score sebesar 91% pada domain ulasan hotel. Namun, penelitian tersebut masih berfokus pada evaluasi model pada domain hotel dan belum banyak mengkaji penerapannya pada domain restoran lokal yang memiliki karakteristik linguistik berbeda, seperti penggunaan bahasa informal, slang, serta keberadaan beberapa opini terhadap aspek yang berbeda dalam satu ulasan. Selain itu, sebagian besar penelitian ABSA masih berfokus pada evaluasi performa model menggunakan metrik-metrik evaluasi. Meskipun metrik evaluasi penting untuk menilai kemampuan model, hasil analisis yang disajikan dalam bentuk nilai

evaluasi teknis sering kali kurang mudah dipahami oleh pengguna non-teknis dan pengambil keputusan bisnis (Firdaus et al., 2024). Di sisi lain, visualisasi data dan dashboard analitik telah terbukti mampu meningkatkan pemahaman terhadap hasil analisis serta memudahkan proses pemahaman informasi oleh pengguna yang tidak memiliki keahlian teknis khusus (Jabar et al., 2025; Lavallo et al., 2024). Oleh karena itu, diperlukan pendekatan yang tidak hanya mampu mengidentifikasi aspek dan sentimen secara akurat, tetapi juga menyajikan hasil analisis melalui visualisasi yang informatif sehingga dapat dimanfaatkan secara lebih efektif dalam proses evaluasi layanan dan pengambilan keputusan.

BPK Tesalonika merupakan salah satu usaha kuliner yang memperoleh ulasan pelanggan dalam jumlah besar dan terus bertambah pada platform Google Maps. Ulasan tersebut memuat berbagai opini terhadap aspek yang berbeda seperti kualitas rasa makanan, pelayanan, harga, suasana, dan lokasi yang dapat dimanfaatkan sebagai sumber evaluasi operasional. Akan tetapi, banyaknya data ulasan menyulitkan proses identifikasi aspek yang paling sering mendapatkan pujian maupun keluhan dari pelanggan apabila dilakukan secara manual. Oleh karena itu penelitian ini menerapkan IndoBERT dengan pendekatan *Sentence Pair Classification* untuk melakukan deteksi aspek dan klasifikasi sentimen pada ulasan pelanggan BPK Tesalonika. Selain mengevaluasi efektivitas pendekatan IndoBERT-SPC pada domain restoran lokal berbahasa Indonesia, penelitian ini juga mengembangkan dashboard

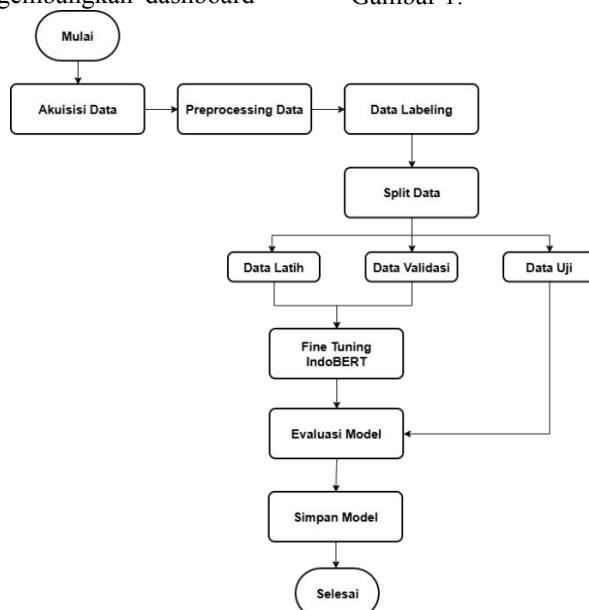
analitik yang menyajikan hasil sentimen berbasis aspek dalam bentuk visualisasi interaktif sehingga dapat membantu pihak manajemen memahami persepsi pelanggan dan mendukung pengambilan keputusan berbasis data.

2. Metode Penelitian

2.1 Tahapan Penelitian

Penelitian ini dilaksanakan melalui serangkaian tahapan komputasional dan rekayasa perangkat lunak yang sistematis. Secara garis besar, tahapan penelitian terbagi menjadi dua fase utama, yaitu fase pengembangan model analisis sentimen dan fase integrasi sistem *dashboard*. Pada fase pertama (pengembangan model), proses diawali dengan akuisisi data ulasan pelanggan dari platform Google Maps menggunakan teknik *web scraping* melalui layanan otomatisasi Apify. Data mentah yang diperoleh kemudian dilakukan pemrosesan seperti pembersihan teks (*cleaning*), normalisasi bahasa tidak baku, dan penyaringan (*filtering*) untuk mengeliminasi gangguan sintaksis.

Setelah struktur teks menjadi seragam, dilakukan pelabelan data dan transformasi format ke dalam arsitektur *Sentence Pair Classification* (SPC). Model komputasi yang dihasilkan kemudian dievaluasi kinerjanya secara terpisah untuk deteksi aspek dan klasifikasi sentimen, kemudian disimpan dan dipersiapkan untuk diintegrasikan ke dalam dashboard interaktif. Rangkaian kerja pada fase pengembangan model komputasi ini diilustrasikan secara lebih terperinci pada Gambar 1.



Gambar 1. Diagram Alur Perancangan Model

Memasuki fase kedua (integrasi sistem), model IndoBERT yang telah disimpan setelah dievaluasi kinerja diimplementasikan ke dalam prototipe dashboard menggunakan *micro-framework Flask* berbasis kerangka kerja flask. Sistem terdiri atas antarmuka pengguna (*frontend*), modul aplikasi (*backend*), dan basis data MySQL. *Frontend* berfungsi menampilkan antarmuka dan hasil analisis, sedangkan *backend* menangani proses pengambilan ulasan

2.2 Transformasi Data dan Pendekatan Sentence Pair Classification

Setelah proses pelabelan aspek dan sentimen selesai dilakukan, dataset masih berbentuk satu ulasan yang dapat memiliki lebih dari satu pasangan aspek dan sentimen (*one-to-many*). Agar sesuai dengan skema masukan *Sentence Pair Classification* (SPC), setiap ulasan ditransformasikan menjadi pasangan kalimat tunggal (*one-to-one*) dengan memadukan teks ulasan sebagai *context* dan satu kategori aspek sebagai *auxiliary sentence*.

Tabel 1. Transformasi Data dan Pelabelan

No	Ulasan (Context)	Aspek (Auxiliary)	Label Aspek	Label Sentimen
1	makanan ok, tempatnya sih standar, cuma parkir agak susah...	Makanan	1	2
2	makanan ok, tempatnya sih standar, cuma parkir agak susah...	Lokasi	1	0
3	makanan ok, tempatnya sih standar, cuma parkir agak susah...	Suasana	1	1
4	makanan ok, tempatnya sih standar, cuma parkir agak susah...	Pelayanan	0	-1
5	makanan ok, tempatnya sih standar, cuma parkir agak susah...	Harga	0	-1

Berdasarkan Tabel 1, setiap pasangan *context-auxiliary* diberikan dua jenis label, yaitu Label Aspek dan Label Sentimen. Label Aspek bernilai 1 apabila kategori aspek terdapat pada ulasan dan 0 apabila aspek tersebut tidak dibahas. Selanjutnya, Label Sentimen hanya diberikan pada pasangan yang memiliki Label Aspek = 1, dengan nilai 0 untuk sentimen negatif, 1 untuk sentimen netral, dan 2 untuk sentimen positif. Sementara itu, pasangan data dengan Label Aspek = 0 diberikan nilai -1 pada Label Sentimen sebagai penanda bahwa tidak terdapat sentimen yang perlu diklasifikasikan karena aspek tersebut tidak muncul dalam ulasan. Dataset hasil transformasi ini kemudian digunakan sebagai masukan pada proses pelatihan model deteksi aspek dan model klasifikasi sentimen.

2.3 Pemodelan Transfer Learning dan Penanganan Ketidakseimbangan Kelas

Pengembangan model dilakukan dengan menerapkan *transfer learning* pada arsitektur bahasa pra-latih *indobert-base-pl* yang

melalui layanan Apify, menjalankan inferensi model, serta mengelola alur pemrosesan data. Basis data MySQL digunakan untuk menyimpan riwayat hasil analisis. Integrasi ini memungkinkan sistem memproses ulasan pelanggan secara otomatis dan menyajikan hasilnya dalam bentuk metrik, distribusi aspek dan sentimen, tren berbasis waktu, serta visualisasi kata kunci sehingga lebih mudah dipahami oleh pengguna.

Melalui proses ini, setiap ulasan diekspansi menjadi lima pasangan data yang merepresentasikan lima kategori aspek, yaitu makanan, pelayanan, harga, suasana, dan lokasi. Transformasi tersebut memungkinkan proses deteksi aspek dan klasifikasi sentimen dilakukan secara terpisah pada setiap pasangan data sehingga sesuai dengan mekanisme pembelajaran model IndoBERT berbasis SPC (Yulianti et al., 2024). Hasil transformasi data ditunjukkan pada Tabel 1.

memiliki 12 *encoder layer*, 768 *hidden size*, dan 12 *attention heads*. Melalui mekanisme *self-attention*, model dapat memahami konteks kata secara *bidirectional* sehingga efektif digunakan untuk tugas klasifikasi sentimen berbasis aspek (Wilie et al., 2020). Mengingat distribusi ulasan restoran rentan terhadap ketidakseimbangan kelas, di mana kelas mayoritas mendominasi data, penelitian ini menerapkan pendekatan *Cost-Sensitive Learning*. Pendekatan ini diimplementasikan melalui modifikasi fungsi *loss* menjadi *Weighted Cross-Entropy* yang memberikan penalti komputasi lebih besar terhadap kesalahan prediksi pada kelas minoritas (Araf et al., 2024). Secara matematis, fungsi objektif tersebut dirumuskan pada persamaan (1).

$$L = - \sum_{i=1}^C w_i y_i \log(p_i) \quad (1)$$

Di mana C menyatakan jumlah kelas, y_i merupakan label sebenarnya untuk kelas ke- i , p_i merupakan probabilitas prediksi model untuk

kelas ke- i , dan w_i merupakan bobot kelas yang diberikan untuk mengurangi pengaruh dominasi kelas mayoritas selama proses pelatihan model.

2.4 Skenario Eksperimen *Hyperparameter* dan Evaluasi Kinerja

Skenario pelatihan model (*fine-tuning*) dirancang menggunakan optimizer AdamW dengan ukuran *batch* 16 dan panjang maksimum sekuens sebesar 128 token. Untuk memperoleh titik konvergensi yang optimal, penelitian ini melakukan eksperimen perbandingan *learning rate* pada nilai 1e-05, 2e-05, 3e-05, dan 5e-05 dengan batas maksimum pelatihan sebanyak 10 *epoch*.

Proses pelatihan dikendalikan menggunakan mekanisme *early stopping* untuk mencegah terjadinya *overfitting*. Evaluasi kinerja model pada tugas deteksi aspek biner dan klasifikasi sentimen multikelas dilakukan menggunakan metrik *Accuracy*, *Precision*, *Recall*, *F1-Score*, dan *Macro F1-Score* yang dihitung berdasarkan *confusion matrix*. Rumusan matematis dari metrik evaluasi tersebut disajikan pada Persamaan (2) hingga Persamaan (6).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

$$Recall = \frac{TP}{TP + FN} \quad (4)$$

$$F1-Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (5)$$

$$Macro F1-Score = \frac{1}{C} \sum_{i=1}^C F1_i \quad (6)$$

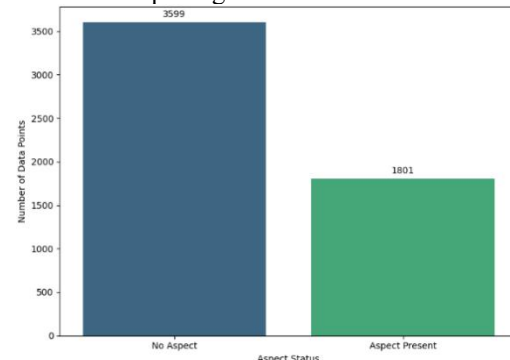
Di mana *True Positive* (TP) menyatakan jumlah data positif yang berhasil diprediksi dengan benar, *True Negative* (TN) menyatakan jumlah data negatif yang berhasil diprediksi dengan benar, *False Positive* (FP) menyatakan jumlah data negatif yang salah diprediksi sebagai positif, dan *False Negative* (FN) menyatakan jumlah data positif yang salah diprediksi sebagai negatif. Selanjutnya, C menyatakan jumlah kelas dan $F1_i$ merupakan nilai *F1-Score* pada kelas ke- i . Penggunaan *Macro F1-Score* bertujuan memberikan bobot evaluasi yang sama pada setiap kelas sehingga

performa model pada kelas mayoritas maupun minoritas dapat diukur secara lebih seimbang (Grandini et al., 2020).

3. Hasil dan Pembahasan

3.1 Distribusi Data dan Ketidakseimbangan Kelas

Sebelum tahapan pelatihan model komputasi dilakukan, eksplorasi terhadap distribusi himpunan data ulasan pelanggan restoran BPK Tesalonika menjadi krusial untuk menentukan strategi pemodelan yang tepat. Setelah dilakukan pemrosesan data dan pelabelan data, diperoleh data ulasan final yang berjumlah 1.080 observasi, lalu ditransformasikan ke format menggunakan pendekatan *Sentence Pair Classification* (SPC), di mana setiap teks ulasan dipasangkan dengan lima kategori aspek, menghasilkan total 5.400 observasi berpasangan.



Gambar 2. Grafik Distribusi Data Deteksi Aspek Biner

Berdasarkan hasil pelabelan data yang divisualisasikan pada Gambar 2, terlihat bahwa dominasi data berada pada kelas ketiadaan aspek (*No Aspect*) sebesar 66,65% (3.599 observasi), sementara ulasan yang benar-benar memuat opini spesifik terhadap suatu aspek (*Aspect Present*) hanya sebesar 33,35% (1.801 observasi). Pada observasi yang tidak memiliki aspek, nilai sentimennya ditetapkan menjadi -1 agar diabaikan dalam kalkulasi kerugian (*loss*) saat pelatihan klasifikasi sentimen.

Selanjutnya, analisis difokuskan pada 1.801 observasi yang relevan untuk melihat distribusi polaritas sentimen terhadap masing-masing kategori aspek operasional, yang disajikan pada Tabel 2.

Tabel 2 Distribusi Label Polaritas Sentimen Per Kategori Aspek

Kategori Aspek	Sentimen Negatif (0)	Sentimen Netral (1)	Sentimen Positif (2)	Total Observasi
Harga	49	29	152	230
Lokasi	44	28	86	158
Makanan	149	46	660	855

Pelayanan	94	34	154	282
Suasana	84	32	160	276
Total	420	169	1.212	1.801

Fakta kuantitatif pada Tabel 2 menunjukkan dua temuan utama mengenai karakteristik ulasan pelanggan BPK Tesalonika. Pertama, distribusi label sentimen mengalami ketidakseimbangan (*class imbalance*), di mana sentimen positif mendominasi sebesar 67,30%, diikuti sentimen negatif sebesar 23,32%, sedangkan sentimen netral hanya 9,38%. Kedua, berdasarkan aspek yang dibahas, makanan menjadi aspek yang paling banyak dievaluasi dengan total 855 observasi. Dominasi sentimen positif pada aspek makanan, yaitu sebanyak 660 observasi, menunjukkan bahwa kualitas rasa dan keaslian hidangan merupakan keunggulan utama BPK Tesalonika. Temuan ini sejalan dengan karakteristik BPK Tesalonika yang dikenal dengan menu utama Babi Panggang Karo (BPK) di mana cita rasa menjadi faktor utama dalam menentukan kepuasan pelanggan. Sebaliknya, aspek lain seperti lokasi hanya memperoleh 158 observasi, yang menunjukkan bahwa pelanggan lebih banyak memberikan perhatian terhadap kualitas makanan dibandingkan faktor lainnya.

Meskipun tingginya sentimen positif mencerminkan keberhasilan restoran dalam memenuhi harapan pelanggan, kondisi ini menjadi tantangan dalam proses pelatihan

model *machine learning*. Ketidakseimbangan distribusi sentimen dapat menyebabkan model lebih cenderung memprediksi kelas mayoritas sehingga performa prediksi pada kelas minoritas menjadi menurun. Oleh karena itu, penelitian ini menerapkan pendekatan *Cost-Sensitive Learning* melalui fungsi *Weighted Cross-Entropy* pada model IndoBERT. Metode ini memberikan bobot yang lebih besar pada kelas minoritas sehingga model dapat mempelajari seluruh kelas sentimen secara lebih seimbang dan mengurangi bias terhadap kelas positif selama proses *fine-tuning*.

3.2 Eksperimen Pelatihan dan Penentuan Model Terbaik

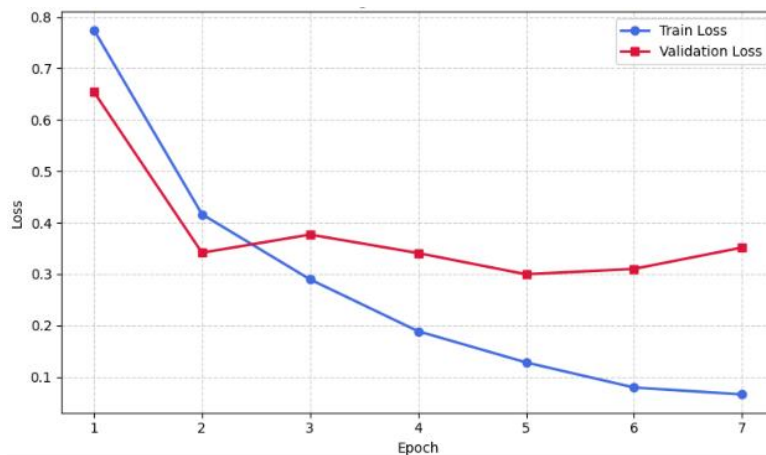
Proses *fine-tuning* model *indobert-base-pl* dilaksanakan di atas platform Google Colab dengan dukungan *Graphics Processing Unit* (GPU) NVIDIA Tesla T4. Skenario pelatihan dieksekusi dengan menguji empat variasi nilai kecepatan pembelajaran (*learning rate*), yakni 1e-05, 2e-05, 3e-05, dan 5e-05 menggunakan fungsi optimasi AdamW. Untuk menjaga model dari risiko kehilangan kemampuan generalisasi pada data baru, diterapkan mekanisme *early stopping* dengan tingkat toleransi dua epoch dari batas maksimal 10 epoch. Hasil komparasi dari eksperimen komputasi tersebut disajikan pada Tabel 3.

Tabel 3. Hasil Eksperimen *Fine-Tuning* Model IndoBERT

Learning Rate	Epoch Berhenti (Titik Optimal)	Validation Loss	Aspect F1-Score	Sentiment Macro F1-Score
5e-05	7 (Optimal di Epoch 5)	0,2994	0,9171	0,8228
1e-05	6 (Optimal di Epoch 4)	0,3210	0,9159	0,8205
2e-05	6 (Optimal di Epoch 4)	0,2822	0,9333	0,8084
3e-05	8 (Optimal di Epoch 6)	0,3039	0,9304	0,7864

Berdasarkan data pada Tabel 3, konfigurasi dengan *learning rate* 2e-05 secara teknis berhasil mencapai tingkat *Validation Loss* terendah (0,2822) serta menghasilkan *Aspect F1-Score* tertinggi (0,9333). Walaupun demikian, skenario tersebut tidak dipilih sebagai model final, karena karakteristik tugas klasifikasi sentimen yang berhadapan dengan masalah ketidakseimbangan kelas membuat

nilai *Macro F1-Score* menjadi indikator keberhasilan yang jauh lebih valid, karena metrik ini menghitung rata-rata performa dengan memperlakukan seluruh kelas secara setara tanpa didominasi oleh sentimen positif. Berdasarkan hal tersebut, eksperimen dengan *learning rate* 5e-05 terbukti memberikan performa yang paling superior dengan nilai *Sentiment Macro F1-Score* tertinggi mencapai 0,8228 pada epoch kelima.



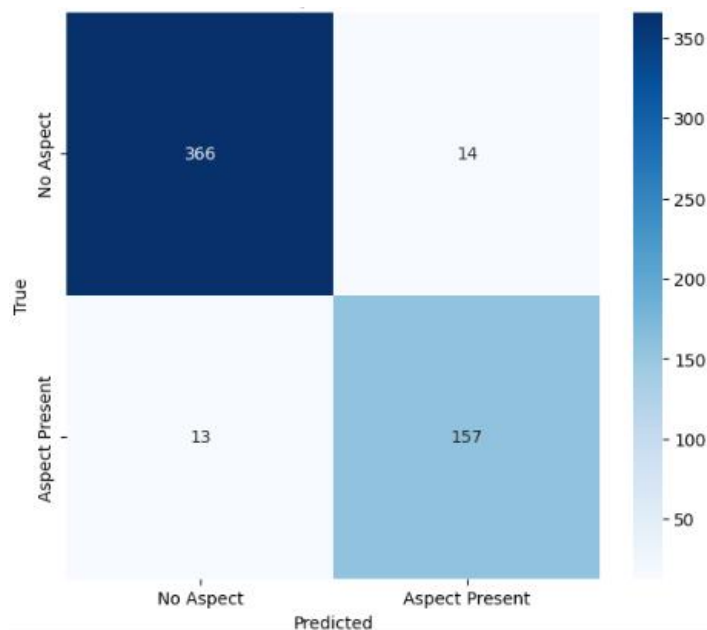
Gambar 3. Grafik Konvergensi Pelatihan Model IndoBERT

Dinamika pergerakan pelatihan model final dengan *learning rate* 5e-05 divisualisasikan pada Gambar 3. Kurva *Training Loss* terlihat menurun secara konsisten dari angka 0,7747 pada epoch pertama hingga menyentuh 0,0657 di epoch ketujuh, merepresentasikan kapasitas model yang semakin baik dalam memetakan pola data latih. Di sisi lain, kurva *Validation Loss* menampilkan karakteristik yang sedikit berfluktuasi dan mencapai rekor terendahnya (0,2994) bertepatan pada epoch kelima. Namun, ketika iterasi dipaksakan berlanjut ke epoch keenam dan ketujuh, kurva *Validation Loss* justru berbalik arah dan mengalami peningkatan hingga mencapai 0,3515. Kenaikan *loss* validasi ini berurutan yang bersifat hierarkis, yakni akurasi deteksi kategori aspek biner dan konsistensi klasifikasi sentimen multikelas

mengindikasikan gejala awal *overfitting*, sehingga mekanisme *early stopping* secara otomatis menghentikan proses pelatihan dan memulihkan bobot parameter terbaik pada epoch kelima untuk dijadikan sebagai model produksi.

3.3 Evaluasi Kinerja Model Analisis Sentimen Berbasis Aspek

Evaluasi kapabilitas generalisasi model dilakukan menggunakan himpunan data uji independen yang terdiri atas 550 observasi *unseen data*. Mengingat sistem ini dibangun menggunakan arsitektur *Decoupled Classifier* atau pemisahan lapisan klasifikasi, maka pengukuran kinerja dibedah menjadi dua tahap.



Gambar 4. Matriks Konfusi Deteksi Aspek Biner

Pada tahap pertama, model dievaluasi kemampuannya dalam mengidentifikasi keberadaan aspek di dalam suatu kalimat ulasan. Hasil uji matriks konfusi pada Gambar 4 menunjukkan bahwa model mengidentifikasi

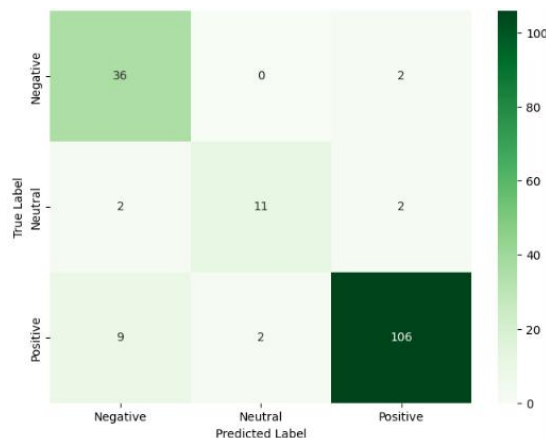
dengan benar 366 observasi *no aspect* (*True Negatives*) dan 157 observasi *aspect present* (*True Positives*) secara presisi. Kesalahan prediksi terdiri atas 14 kejadian *False Positives* dan 13 kejadian *False Negatives*.

Tabel 4. Laporan Klasifikasi Deteksi Aspek Biner

Kelas	Precision	Recall	F1-Score	Support
No Aspect	0.9657	0.9632	0.9644	380
Aspect Present	0.9181	0.9235	0.9208	170
Accuracy	-	-	0.9509	550
Macro Average	0.9419	0.9433	0.9426	550
Weighted Average	0.9510	0.9509	0.9509	550

Berdasarkan Tabel 4, laporan kinerja untuk tugas deteksi aspek memperoleh tingkat Akurasi global sebesar 95,09%. Nilai *F1-Score* pada pengenalan kelas keberadaan aspek mencapai 92,08% dengan tingkat Presisi (91,81%) dan

Recall (92,35%). Hasil ini menunjukkan bahwa mesin pengolah bahasa berbasis *Transformer* menunjukkan kinerja yang baik, dalam mengenali konteks teks terhadap target pilar operasional restoran dengan tingkat kesalahan klasifikasi yang rendah.



Gambar 5. Matriks Konfusi Klasifikasi Sentimen Multikelas

Tahap pengujian kedua berfokus pada kemampuan model membedakan polaritas opini pelanggan menjadi sentimen positif, negatif, atau netral. Sesuai dengan skenario evaluasi, evaluasi ini hanya melibatkan 170 data uji yang sebelumnya telah diidentifikasi memiliki aspek.

Matriks konfusi pada Gambar 5 memperlihatkan kinerja klasifikasi yang baik, di mana model mengklasifikasikan dengan benar 106 observasi sentimen positif, 36 observasi negatif, dan 11 observasi netral.

Tabel 5. Laporan Klasifikasi Sentimen Multikelas

Kelas	Precision	Recall	F1-Score	Support
Negative	0.7660	0.9474	0.8471	38
Neutral	0.8462	0.7333	0.7857	15
Positive	0.9636	0.9060	0.9339	117
Accuracy	-	-	0.9000	170
Macro Average	0.8586	0.8622	0.8556	170
Weighted Average	0.9091	0.9000	0.9014	170

Merujuk pada Tabel 5, performa pada tahap klasifikasi sentimen ini mencapai Akurasi global sebesar 90,00% dengan nilai *Macro F1-Score* sebesar 85,56%. Jika ditinjau per kelas, nilai F1-Score tertinggi diperoleh oleh sentimen

Positif dengan nilai *F1-Score* mencapai 93,39%, disusul Negatif sebesar 84,71%, dan Netral sebesar 78,57%. Penurunan margin performa pada kelas sentimen netral ini merupakan fenomena yang dapat dijelaskan dan

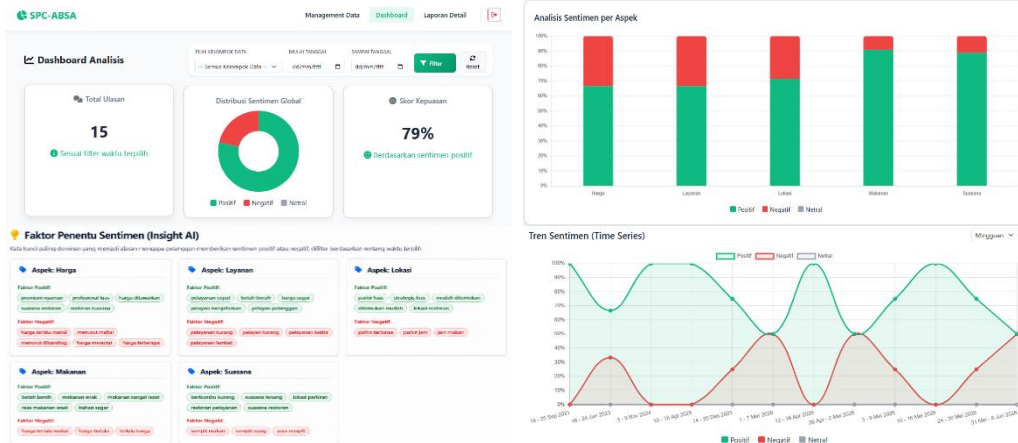
dipengaruhi secara langsung oleh keterbatasan jumlah sampel pelatihan (hanya 9,38%) pada kelas tersebut. Walaupun demikian, angka 78,57% menunjukkan bahwa intervensi

3.4 Implementasi dan Visualisasi pada Dashboard Interaktif

Hasil prediksi model IndoBERT diintegrasikan ke dalam dashboard berbasis Python Flask. Model yang telah dievaluasi disimpan dalam format PyTorch dan

pembobotan *Weighted Cross-Entropy* membantu meningkatkan mengurangi kecenderungan model dari bias klasifikasi yang umumnya merugikan kelas minoritas.

diimplementasikan menggunakan struktur *Model View Controller Service* (MVCS), integrasi ini memungkinkan proses prediksi dijalankan pada server lokal menggunakan CPU tanpa memerlukan GPU.



Gambar 6. Rangkaian Visualisasi Dashboard Analisis Interaktif

Hasil integrasi sistem ditampilkan yang di tampilkan pada Gambar 6, dashboard menyajikan hasil analisis sentimen dalam bentuk visual untuk mendukung pemantauan ulasan pelanggan. Dashboard ini merangkum data ulasan ke dalam empat komponen utama. Pertama, kartu ringkasan menampilkan jumlah ulasan yang dianalisis, distribusi sentimen secara keseluruhan, dan persentase tingkat kepuasan pelanggan. Kedua, grafik analisis sentimen per aspek menampilkan distribusi sentimen positif, negatif, dan netral pada lima aspek, yaitu makanan, pelayanan, harga, suasana, dan lokasi. Ketiga, modul *Gradient Saliency* digunakan sebagai metode *Explainable AI* untuk menyoroti kata-kata yang memberikan kontribusi terhadap prediksi sentimen model. Keempat, grafik tren sentimen menyajikan perubahan distribusi sentimen berdasarkan waktu sehingga memudahkan pemantauan perkembangan kepuasan pelanggan. Melalui dashboard ini, hasil analisis sentimen dapat dipantau secara lebih terstruktur sehingga membantu pemilik restoran dalam mengidentifikasi aspek yang memperoleh respons positif maupun negatif dari pelanggan.

4. Kesimpulan

Penelitian ini mengimplementasikan analisis sentimen berbasis aspek pada ulasan pelanggan restoran BPK Tesalonika

menggunakan model IndoBERT dengan pendekatan *Sentence Pair Classification* (SPC). Berdasarkan hasil eksperimen, konfigurasi *learning rate* 5e-05 memberikan performa terbaik. Pada data uji, model mencapai akurasi 95,09% dan F1-Score 92,08% untuk tugas deteksi aspek biner, serta akurasi 90,00% dan Macro F1-Score 85,56% untuk klasifikasi sentimen multikelas. Hasil tersebut menunjukkan bahwa model mampu mengidentifikasi aspek dan mengklasifikasikan sentimen pelanggan dengan baik. Namun, performa pada kelas sentimen netral masih lebih rendah dibandingkan kelas lainnya karena jumlah data latih yang terbatas.

Model yang telah dikembangkan kemudian diintegrasikan ke dalam prototipe dashboard berbasis Flask untuk menyajikan hasil analisis dalam bentuk visual. Dashboard menampilkan ringkasan sentimen, distribusi sentimen pada lima aspek utama (makanan, pelayanan, harga, suasana, dan lokasi), tren sentimen berdasarkan waktu, serta kata kunci yang berkontribusi terhadap prediksi menggunakan *Gradient Saliency*. Visualisasi ini membantu pengguna memahami hasil analisis sentimen secara lebih mudah dan terstruktur. Penelitian selanjutnya disarankan meningkatkan jumlah data dari restoran dengan menu dan layanan serupa untuk menyeimbangkan distribusi kelas. Selain itu, aspek dapat dikembangkan menjadi subaspek yang lebih spesifik untuk meningkatkan

ketepatan analisis. Selain itu penelitian selanjutnya disarankan menerapkan pendekatan *token-level classification* agar kata atau frasa yang memengaruhi prediksi sentimen dapat diidentifikasi secara lebih rinci.

REFERENCES

- [1] I. Araf, A. Idri, and I. Chairi, *Cost - sensitive learning for imbalanced medical data : a review*, vol. 57, no. 4. Springer Netherlands, 2024. doi: 10.1007/s10462-023-10652-8.
- [2] L. Aureliano-silva and E. E. Spers, "The effect of online reviews on restaurant visit intentions : applying signaling and involvement theories," *Eff. online Rev. Restaur. Visit intentions Appl. Signal. Invol. Theor. Leonardo*, vol. 12, no. 4, pp. 672–688, 2021, doi: 10.1108/JHTT-06-2020-0143.
- [3] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," *NAACL HLT 2019 - 2019 Conf. North Am. Chapter Assoc. Comput. Linguist. Hum. Lang. Technol. - Proc. Conf.*, vol. 1, no. Mlm, pp. 4171–4186, 2019.
- [4] M. F. Firdaus, D. E. Ratnawati, and N. Y. Setiawan, "ANALISIS SENTIMEN BERBASIS ASPEK PADA ULASAN PELANGGAN RESTORAN MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE (STUDI KASUS : DEPOT BAMARA) ASPECT-BASED SENTIMENT ANALYSIS ON RESTAURANT CUSTOMER REVIEWS USING SUPPORT VECTOR MACHINE ALGORITHM (CASE STUDY," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 11, no. 6, pp. 1265–1272, 2024, doi: 10.25126/jtiik.2024117564.
- [5] E. Bagli and G. Visani, "M m -c c ;," *METRICS MULTI-CLASS Classif. AN Overv.*, pp. 1–17, 2020.
- [6] Y. Cathy, H. Paul, D. Jörg, and W. Katerina, *A systematic review of aspect - based sentiment analysis : domains , methods , and trends*, vol. 57, no. 11. Springer Netherlands, 2024. doi: 10.1007/s10462-024-10906-z.
- [7] A. A. Jabar, R. F. Wijaya, S. Wahyuni, U. Pembangunan, P. Budi, and A. Info, "DASHBOARD VISUALISASI DATA KECERDASAN BISNIS," *Ami Abdul Jabar1, Rian Farta Wijaya2, Sri Wahyuni3*, vol. 6, no. 1, pp. 75–89, 2025, doi: 10.46576/djtechno.
- [8] A. Lavallo, A. Mat, J. Trujillo, and S. Rizzi, "Visualization Requirements for Business Intelligence Analytics : A Goal-Based , Iterative Framework," *Vis. Requir. Bus. Intell. Anal. A Goal-Based, Iterative Framew. Ana*, 2024.
- [9] "Restaurant survival prediction using customer-generated content: An aspect-based sentiment analysis of online reviews Hengyun Li," *Restaur. Surviv. Predict. using Cust. content An Asp. Sentim. Anal. online Rev.*, pp. 1–33, 2023, doi: 10.1016/j.tourman.2022.104707.
- [10] S. Li, F. Liu, Y. Zhang, B. Zhu, and H. Zhu, "Text Mining of User-Generated Content (UGC) for Business Applications in E-Commerce : A Systematic Review," *Text Min. User-Generated Content Bus. Appl. E-Commerce A Syst. Rev.*, pp. 1–26, 2022, doi: 10.3390.
- [11] B. Liu, "Sentiment Analysis and Opinion Mining," *Sentim. Anal. Opin. Min.*, no. May, pp. 1–168, 2012.
- [12] C. Lohith and H. Chandramouli, "Fine-Grained Sentiment Analysis Of Restaurant Reviews Using Latent Dirichlet Allocation And Hybrid Ensembles," *Int. J. Environ. Sci.*, vol. 11, no. 20, pp. 2468–2475, 2025.
- [13] M. Pontiki and J. Pavlopoulos, "SemEval-2014 Task 4 : Aspect Based Sentiment Analysis," *Asp. Based Sentim. Anal.*, no. SemEval, pp. 27–35, 2014.
- [14] C. Sun, L. Huang, and X. Qiu, "Utilizing BERT for Aspect-Based Sentiment Analysis via Constructing Auxiliary Sentence," *Util. BERT Asp. Sentim. Anal. via Constr. Aux. Sentence Chi*, pp. 380–385, 2019.
- [15] A. Vaswani, "Attention Is All You Need," *Atten. Is All You Need*, no. Nips, 2017.
- [16] B. Wilie *et al.*, "IndoNLU : Benchmark and Resources for Evaluating Indonesian Natural Language Understanding," *IndoNLU Benchmark*

- Resour. Eval. Indones. Nat. Lang. Underst. Bryan*, pp. 843–857, 2020.
- [17] E. Yulianti and N. K. Nissa, “ABSA of Indonesian customer reviews using IndoBERT: single-sentence and sentence-pair classification approaches,” *Bull. Electr. Eng. Informatics*, vol. 13, no. 5, pp. 579589, 2024, doi: 10.11591/eei.v13i5.80