



Identifikasi Tingkat Intensitas Opini dalam Analisis Sentimen Berbasis Aspek Menggunakan Enhanced Triplet Extraction

Gabriel Jimmy Richardo Chastelo B¹, Sunneng Sandino Berutu^{2,*}, Heani Budiati¹

¹ Fakultas Sains dan Komputer, Program Studi Informatika, Universitas Kristen Immanuel, Yogyakarta, Indonesia

² Fakultas Sains dan Komputer, Program Studi Informatika, Universitas Duta Bangsa Surakarta, Surakarta, Indonesia

Email: ¹gabriel.jimmy.r@mail.ukrim.ac.id, ^{2,*}sandinoberutu@udb.ac.id, ³heni@ukrimuniversity.ac.id

Email Penulis Korespondensi: sandinoberutu@udb.ac.id

Abstrak—Dalam analisis sentimen konvensional sering kali mengabaikan variasi tingkat intensitas opini dalam ulasan teks. Hal ini disebabkan oleh keterbatasan *task* pada pendekatan *Aspect-Based Sentiment Analysis* (ABSA) yang dibatasi pada tiga komponen *triplet* utama. Penelitian ini bertujuan untuk mengembangkan dan memperluas kerangka kerja *Aspect-Sentiment-Opinion Triplet Extraction* (ASOTE) dalam melakukan ekstraksi hubungan antar-entitas dan polaritas sentimen dengan mengintegrasikan deteksi intensitas opini. Penelitian ini mengimplementasikan pendekatan ABSA dengan memperluas struktur *triplet* menjadi empat komponen, yaitu aspek, opini, intensifier, dan sentimen (*Enhanced Triplet*). Data dikumpulkan melalui proses *web scraping* pada komentar Twitter (X) terkait program *Makan Bergizi Gratis* yang digunakan sebagai studi kasus untuk menguji kemampuan model dalam menganalisis sentimen publik. Data kemudian melalui tahap pra-pemrosesan dan pelabelan BIO Tagging, serta diklasifikasikan menggunakan pendekatan *fine-grained sentiment* untuk menangkap nuansa intensitas emosi secara lebih rinci. Model berbasis *Transformer*, yaitu IndoBERT, digunakan untuk memahami konteks dan intensitas makna dalam bahasa Indonesia. Hasil evaluasi pada data uji menunjukkan bahwa model mampu mencapai akurasi sebesar 88% dan nilai rata-rata *F1-Score* sebesar 0.88 dalam klasifikasi polaritas sentimen antar entitas, yang menunjukkan performa model yang baik. Hasil ini membuktikan dengan menyediakan kerangka kerja yang lebih peka terhadap intensitas opini dalam mengklasifikasikan nuansa emosi publik merupakan solusi yang sangat efektif.

Kata Kunci: ABSA; ASOTE; IndoBERT; Intensitas Opini; BIO Tagging; Fine-grained Sentiment

Abstract—Conventional sentiment analysis often overlooks variations in the intensity of opinions within text reviews. This is due to the limitations of the Aspect-Based Sentiment Analysis (ABSA) approach, which is restricted to three main triplet components. This study aims to develop and expand the Aspect-Sentiment-Opinion Triplet Extraction (ASOTE) framework to extract entity relationships and sentiment polarity by integrating opinion intensity detection. This study implements the ABSA approach by expanding the triplet structure into four components: aspect, opinion, intensifier, and sentiment (Enhanced Triplet). Data was collected via web scraping of Twitter (X) comments related to the Free Nutritious Meals program, which served as a case study to test the model's ability to analyze public sentiment. The data then undergoes pre-processing and BIO Tagging, and is classified using a fine-grained sentiment approach to capture the nuances of emotional intensity in greater detail. A Transformer-based model, namely IndoBERT, was used to understand the context and intensity of meaning in the Indonesian language. Evaluation results on the test data show that the model achieved an accuracy of 88% and an average F1-score of 0.88 in sentiment polarity classification between entities, indicating strong model performance. These results demonstrate that providing a framework that is more sensitive to the intensity of opinions when classifying the nuances of public sentiment is a highly effective solution.

Keywords: ABSA; ASOTE; IndoBERT; Opinion Intensity; BIO Tagging; Fine-grained Sentiment

1. PENDAHULUAN

Aspect-Based Sentiment Analysis (ABSA) merupakan suatu metode mutakhir dalam bidang pemrosesan bahasa alami/*Natural Language Processing* (NLP) yang dirancang secara spesifik untuk menganalisis dan menginterpretasikan sentimen pengguna dengan tingkat ketelitian dan presisi yang lebih tinggi[1][2][3]. Berbeda dengan metode analisis sentimen konvensional yang hanya menghasilkan penilaian umum mengenai polaritas sentimen sentimen umum dengan mengklasifikasikan teks sebagai positif, negatif, dan netral. ABSA meningkatkan kemampuan analisis secara sistematis mendekonstruksi sentimen menjadi komponen atau aspek berbeda. Hal ini didasarkan pada fakta bahwa dalam satu kalimat ulasan, pengguna dapat mengekspresikan opini yang berbeda terhadap berbagai aspek yang terkandung di dalamnya. Oleh karena itu implementasi ABSA tidak hanya memperkaya pemahaman terhadap suatu opini tetapi juga memfasilitasi evaluasi setiap aspek secara lebih komprehensif sehingga pada akhirnya menghasilkan analisis yang lebih mendalam dan akurat[4].

Dalam ranah analisis sentimen tradisional pernyataan seperti “Kamera ponsel ini sangat jelas, tetapi baterainya sangat boros”. sering diklasifikasikan sebagai sentimen netral atau campuran oleh sistem otomatis yang sering menyebabkan ambigu. Pada titik ini ABSA memainkan peran penting, metode ini akan menguraikan entitas tersebut menjadi dua komponen terpisah yakni aspek “Kamera” dengan polaritas positif serta aspek “Baterai” dengan polaritas negatif. Melalui pendekatan ini dapat memperoleh data yang lebih presisi terkait kekuatan dan kelemahan analisis sentimen. Dengan menggunakan pendekatan ini dapat diperoleh informasi yang lebih presisi mengenai aspek-aspek spesifik yang dievaluasi dalam suatu ulasan atau komentar sehingga menghasilkan analisis yang lebih komprehensif dibandingkan pendekatan klasifikasi tradisional yang hanya menghasilkan satu label sentimen secara keseluruhan.

Pada pendekatan ABSA terdapat tiga komponen aspek, opini dan sentimen yang membentuk suatu sistem saling berkaitan. Untuk menjalankan tiga komponen bersamaan pendekatan ABSA sangat bergantung pada model kecerdasan buatan berbasis *deep learning*, khususnya arsitektur *Transformer* seperti BERT[5][6]. Terdapat tiga tugas utama yang harus diselesaikan secara bersamaan yakni ekstraksi aspek untuk mengidentifikasi entitas dan atribut, deteksi opini untuk menemukan kata penilaian dari aspek, dan klasifikasi sentimen untuk menentukan polaritas sentimen. Berbeda dengan



pendekatan tradisional yang membangun komponen secara terpisah, deep learning memungkinkan ketiga tugas ini dipelajari secara bersamaan dalam satu model[7][8].

Untuk mencapai tingkat akurasi yang mendalam ABSA melibatkan serangkaian sub-tugas (task) yang memiliki fungsi spesifik dalam menguraikan setiap elemen bahasa. Sub-tugas tersebut mencakup identifikasi entitas hingga ekstraksi hubungan sentimen yang kompleks. Dimulai dari *Aspect Term Extraction* (ATE) untuk menemukan kata/frasa yang merepresentasikan aspek dalam sebuah kalimat. Pada dasarnya ATE berperan sebagai mekanisme identifikasi yang memiliki kapabilitas untuk mengidentifikasi elemen spesifik yang menjadi objek evaluasi dalam suatu teks[9]. Setelah aspek berhasil diidentifikasi tasks berikutnya adalah *Aspect Term Sentiment Analysis* (ATSA). Merupakan suatu tahapan yang berfungsi untuk mengidentifikasi polaritas sentimen baik positif, negatif, maupun netral terhadap aspek tertentu dalam menangkap nuansa opini sebuah teks[10]. Adapun sub-tugas yang lebih spesifik yaitu *Target-oriented Opinion Words Extraction* (TOWE) memiliki fungsi untuk mengekstraksi kata-kata opini dengan aspek spesifik dalam suatu kalimat dan berperan penting dalam analisis sentimen berbasis aspek terutama dalam mengidentifikasi hubungan antara kata opini dan aspek yang relevan[11]. Meskipun TOWE mampu memetakan hubungan antara target dan opini, efektivitas analisis ini semakin disempurnakan melalui sub-tugas (tasks) *Aspect Sentiment Triplet Extraction* (ASTE) yang mampu mengekstrak tiga elemen sekaligus yakni aspek, sentimen, dan opini melalui pendekatan ini sistem bukan hanya sekedar mendeteksi apakah ulasan tersebut bagus atau buruk tetapi memahami apa yang dinilai dan mengapa penilaian tersebut diberikan[12].

Pada penelitian awal dalam ABSA umumnya berfokus pengenalan serta pemetaan hubungan dasar antar aspek, opini, dan sentimen. Salah satu pendekatan yang dikembangkan adalah penerapan model *Large Language Model* (LLM) melalui implementasi arsitektur *Duality-Drive* ASTE yang bertujuan menyempurnakan interaksi antara aspek, opini, dan sentimen guna mengurangi kesalahan ekstraksi. Meskipun pendekatan ini terbukti efektif dalam meningkatkan akurasi ekstraksi triplet konvensional, model tersebut masih terbatas pada pemodelan hubungan tiga komponen dasar (aspek-opini-sentimen) tanpa mempertimbangkan kompleksitas relasi yang lebih luas[13]. Selain pendekatan berbasis LLM, penelitian lain juga berfokus pada pemodelan *transformer* untuk memetakan antara aspek, opini dan polaritas sentimen agar proses penentuan label klasifikasi tidak dilakukan secara terpisah[14]. Perkembangan terbaru dalam ekstraksi triplet telah mengarahkan penerapan model Multi-view Contrastive Learning (MCL) yang bertujuan untuk mengidentifikasi istilah aspek, istilah opini, dan polaritas sentimen untuk membedakan elemen yang benar dan salah secara lebih tegas pada akurasi tiga komponen dasar[15].

Adapun metode alternatif dalam ekstraksi triplet dikembangkan melalui model Deep Relationship Enhancement Networks (DREN) penelitian tersebut memanfaatkan *fitur part of speech* (POS), sintaksis, dan jarak struktural untuk memperkaya pemetaan relasi aspek dan opini dalam satu langkah ekstraksi[16]. Selain pemetaan relasi aspek dan opini ada juga penelitian yang mengembangkan strategi pre-training berbasis contrastive learning yang dikombinasikan dengan pendekatan multi-task learning untuk meningkatkan ketepatan dalam mengidentifikasi aspek, sentimen dan opini[17]. Upaya lain untuk meningkatkan akurasi ekstraksi triple dan mengatasi keterbatasan model sebelumnya penelitian tersebut mencoba memanfaatkan basis pengetahuan *SenticNet* untuk memperbaiki hubungan antara elemen triplet dari setiap kata dan menggabungkannya dengan grafik dependensi Graph Convolutional Network (GCN)[18]. Perkembangan terkini dalam mengeksplorasi penggunaan kerangka ekstraksi dua arah dan strategi *sequential prompting* ini mengintegrasikan hasil ekstraksi tahap sebelumnya sebagai basis pengetahuan untuk mengarahkan proses prediksi target pada tahap berikutnya selain itu penerapan mekanisme dua fungsi kerugian untuk mengoptimalkan interaksi antara triplet aspek, opini, dan sentimen[19]. Untuk menyempurnakan pendekatan sub-tugas sebelumnya, *Aspect-Sentiment-Opinion Triplet Extraction* (ASOTE) dikembangkan sebagai metode lanjutan yang mengalami perkembangan signifikan. Berbeda dengan ASTE yang mengekstrak ketiga elemen secara sekaligus. ASOTE menerapkan pendekatan bertahap melalui model terintegrasi yang mencakup empat komponen utama yaitu ekstraksi aspek, ekstraksi opini, menghubungkan aspek ke opini yang relevan, serta menentukan polaritas sentimen (positif/negatif/netral) dari pasangan aspect-opinion tersebut[20]. Seiring dengan kemajuan yang dicapai dalam berbagai sub-tugas *Aspect-Based Sentiment Analysis* (ABSA) mulai dari identifikasi aspek hingga ekstraksi triplet telah menunjukkan keberhasilan dalam mengekstraksi aspek, opini, dan polaritas sentimen dalam teks ulasan. Meskipun telah menunjukkan keberhasilan dalam mengekstraksi ketiga triplet aspek, opini, sentimen pendekatan ini memiliki keterbatasan signifikan dalam menangkap intensitas opini yang diekspresikan melalui penegasan makna (*intensifier*). Dalam penyampaian argumen tidak semua mengekspresikan penilaian yang sama seperti “produk bagus” dan “produk bagus banget” dari perbedaan nuansa emosional tersebut belum dapat diakomodasi oleh pendekatan yang ada. Hingga saat ini belum ada penelitian yang mengeksplorasi penambahan elemen intensifier ke dalam triplet sub-tugas ABSA maupun mengembangkan sistem klasifikasi yang mampu membedakan nuansa sentimen lebih terperinci.

Berbagai penelitian sebelumnya menunjukkan bahwa perkembangan ekstraksi triplet dalam *Aspect-Based Sentiment Analysis* (ABSA) telah mengalami kemajuan signifikan, khususnya dalam task ASOTE. Namun, sebagian besar penelitian masih berfokus pada struktur triplet konvensional yang mencakup aspek, opini, dan polaritas sentimen. Meskipun beberapa studi dalam *fine-grained sentiment analysis* telah mempertimbangkan intensitas sentimen, integrasi *intensifier* sebagai komponen eksplisit dalam struktur triplet masih terbatas, terutama dalam konteks bahasa Indonesia. Hal ini menunjukkan adanya celah penelitian dalam menggabungkan intensitas ke dalam kerangka ekstraksi triplet. Temuan ini sejalan dengan *Fine-grained ABSA (Transformer-BiLSTM)* yang belum secara eksplisit memodelkan *intensifier* sebagai bagian dari triplet[21]. Oleh karena itu, penelitian ini mengusulkan perluasan struktur triplet menjadi empat komponen (*Enhanced Triplet*) yaitu aspek, opini, *intensifier*, dan sentimen. Pendekatan ini memungkinkan analisis



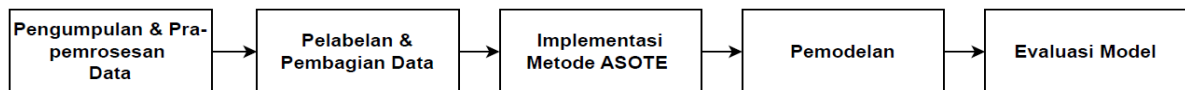
yang lebih mendalam terhadap hubungan antar entitas serta mampu menangkap variasi ekspresi sentimen secara lebih granular dibandingkan pendekatan konvensional. Selain itu, penelitian ini mengusulkan sebuah kerangka kerja berbasis BERT yang sadar posisi entitas *intensifier* untuk menangani task ASOTE. Serta melatih model untuk menemukan entitas dan menentukan polaritas sentimen secara bersamaan dalam memproses bahasa Indonesia.

Penelitian ini memberikan kontribusi pada pengembangan NLP dalam bahasa Indonesia, khususnya dalam deteksi sentimen *fine-grained* yang lebih akurat untuk evaluasi kebijakan publik. Hal ini memberikan perspektif baru dalam pengembangan tugas ABSA dengan menyediakan kerangka kerja yang lebih peka terhadap intensitas opini pada teks informal bahasa Indonesia. Pendekatan ini memberikan pemahaman yang lebih dalam tentang intensitas sentimen pengguna, tidak sekadar terbatas pada polaritas dasar (positif atau negatif).

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Penelitian ini menggunakan pendekatan eksperimental kuantitatif yang berfokus pada identifikasi intensitas opini. Dengan mengembangkan metode *Aspect-Sentiment-Opinion Triplet Extraction* (ASOTE) yang diperluas untuk mengekstraksi empat komponen utama yaitu aspek, opini, intensifier, dan polaritas sentimen. Tahapan penelitian meliputi: pengumpulan data, pra-pemrosesan data, pelabelan data, pembagian data latih dan data uji, implementasi metode ASOTE, dan evaluasi kinerja model. Proses tahapan penelitian ini ditunjukkan pada Gambar 1.



Gambar 1 Tahapan Metodologi Penelitian

2.1.1 Pengumpulan Data dan Pra-pemrosesan Data

Pemilihan platform tersebut dilandasi oleh karakteristik data Twitter yang bersifat *real-time* berfungsi sebagai representasi langsung dari opini publik, terutama dalam mengungkap berbagai tanggapan serta ekspresi atau perasaan yang muncul dalam argumen yang disampaikan oleh pengguna. Pengumpulan data dilakukan melalui proses web scraping dimana sistem diprogram untuk menarik data berdasarkan parameter kata kunci 'Program Makan Bergizi Gratis' dalam rentang waktu 1 Januari hingga 30 Juni 2025. Proses pengumpulan data dari platform media sosial Twitter (X) umumnya bersifat tidak terstruktur serta mengandung berbagai elemen noise yang dapat mempengaruhi kualitas analisis. Ini bertujuan untuk membersihkan dan mengubah teks dalam dataset menjadi format yang dapat dipahami oleh model. Langkah-langkah pra-pemrosesan yang diterapkan dalam penelitian ini meliputi komponen-komponen sebagai berikut:

- Pembersihan Data:** Menghapus elemen-elemen yang tidak memiliki nilai semantik maupun informasi yang relevan bagi proses pemodelan mencakup URL, emoji, tanda baca yang berlebihan dan hashtag
- Deduplikasi dan Penyaringan Bahasa:** Menjaga integritas dataset serta mencegah terjadinya bias pada model akibat keberadaan data yang berulang, dilakukan proses penghapusan data duplikat. Selain itu melakukan penyaringan bahasa untuk memastikan bahwa hanya teks berbahasa Indonesia yang diproses.
- Case Folding:** Untuk menyeragamkan format teks dengan mengonversi seluruh karakter menjadi huruf kecil, sehingga model dapat mengidentifikasi kata yang sama secara konsisten tanpa terpengaruh oleh penggunaan huruf kapital.
- Normalisasi Teks:** Mengubah kata-kata tidak baku menjadi baku, sebagai contoh kata "makanggih" dikoreksi menjadi "makan" atau "bgt" menjadi "banget".
- Tokenisasi (*Tokenization*):** Proses ini memecahkan kalimat menjadi unit-unit yang lebih kecil yang disebut token. Fungsi utama dari proses tokenisasi adalah memisahkan struktur kalimat agar dapat dianalisis secara individual pada tingkat kata. Sebagai contoh, kalimat "pelayanannya cepat banget" dipecah menjadi token ('pelayanannya', 'cepat', 'banget').
- Stopword Removal:** Menghapus kalimat atau kata-kata umum yang tidak termasuk pada makna sentimen seperti (dan, yang, di, ke) tujuannya untuk memfokuskan model pada kata-kata kunci yang membawa informasi aspek atau opini.
- Stemming:** Mengubah kata ke bentuk dasar hal ini bertujuan untuk mengurangi perubahan bentuk kata sehingga model dapat mengenali kesamaan makna antar kata.

2.1.2 Pelabelan Data

Setelah melalui tahap pra-pemrosesan data, dilakukan proses pelabelan untuk membentuk *ground truth* sebagai dasar pelatihan model. Pelabelan ini menggunakan skema *BIO Tagging* untuk mengidentifikasi kata atau frasa yang mengindikasikan objek (aspek), penilaian (opini), dan seberapa kuat penilaian yang disampaikan (*intensifier*) dalam setiap kalimat. Dalam penelitian ini, urutan pemrosesan data disusun secara sistematis, yaitu dimulai dari proses scraping, dilanjutkan dengan pembersihan data (*cleaning*), kemudian pelabelan menggunakan skema *BIO Tagging*. Setelah proses pelabelan selesai, tahap stemming diterapkan secara opsional dan selektif.

- Pelabelan BIO Tagging**



Untuk mengidentifikasi entitas aspek, opini, dan *intensifier* ke dalam format yang dapat dipelajari oleh model, penelitian ini menerap pelabelan BIO (Begin, Inside, Outside) untuk mendefinisikan batasan setiap komponen kata secara terstruktur melalui pemberian label (tag) pada setiap token. Melalui pemetaan ini model dapat mengenali posisi serta ruang lingkup dari setiap entitas yang dicari yakni aspek, opini, dan intensifier. Untuk menentukan pelabelan BIO setiap kata (token) dipetakan ke dalam beberapa kategori label:

1. O (Outside): Untuk kata-kata di luar entitas yang tidak memiliki informasi relevan bagi sistem.
 2. B-ASP (Begin): Menandai bagian awal dari kata Aspek
 3. I-ASP (Inside): Menandai kata lanjutan dari Aspek
 4. B-OPI (Begin): Menandai bagian awal dari kata Opini
 5. I-OPI (Inside): Menandai kata lanjutan dari Opini
 6. B-INT (Begin): Menandai bagian awal dari kata Intensifier
- b. Distribusi Sentimen
- Pada tahap klasifikasi sentimen, penelitian ini menggunakan pendekatan *Aspect-Based Sentiment Analysis* (ABSA) dengan konstruksi input berbasis *konteks-target* (*sentence-pair*). Setiap data direpresentasikan dalam format: [CLS] kalimat utuh, [SEP] aspek-opini-intensifier, sehingga model dapat mempelajari hubungan antara konteks ulasan dan triplet entitas yang telah diekstraksi. Pendekatan ini tidak menggunakan modifikasi arsitektur khusus, melainkan memanfaatkan arsitektur standar IndoBERT yang telah di-*fine-tuning* untuk tugas klasifikasi sentimen berbasis aspek. Ini memungkinkan model untuk memahami peran masing-masing entitas dalam menentukan polaritas sentimen secara kontekstual. Seluruh token kemudian diproses menggunakan model IndoBERT dengan arsitektur *sequence classification*, dimana representasi dari token [CLS] digunakan sebagai representasi keseluruhan input (*pooling*) dan diteruskan ke lapisan klasifikasi untuk menghasilkan probabilitas kelas sentimen ke dalam lima kategori, yaitu Sangat Positif, Positif, Netral, Negatif, dan Sangat Negatif.

2.1.3 Pembagian Data

Untuk memastikan model IndoBERT dapat mengenali pola dengan baik pada data baru dan menghindari kondisi *overfitting* karena model terlalu pintar dalam menghafal data latihan namun gagal memahami pola secara umum. Penelitian ini menerapkan skema pembagian data ke dalam dua sub-set independen. Dataset yang telah memiliki label BIO dan label sentimen dibagi menjadi 80% untuk data latih (Training Set), 10% untuk data validasi (validation set) dan 10% untuk data uji (Testing Set). Pembagian data ini bertujuan untuk melatih model dalam memetakan entitas yang merepresentasikan aspek, opini, dan *intensifier*. Sementara itu, uji digunakan untuk mengevaluasi kemampuan model dalam mengenali pasangan antar-entitas dan menentukan polaritas sentimen.

2.1.4 Implementasi Metode ASOTE

Berbeda dengan ASOTE konvensional yang hanya mengekstraksi triplet aspek, opini, dan sentimen, pendekatan dalam penelitian ini memperluas cakupan dengan menambahkan elemen intensitas opini (*intensifier*), sehingga mampu mengidentifikasi empat komponen secara bersamaan, yakni aspek, opini, intensifier, dan polaritas sentimen. Keempat komponen tersebut diekstraksi secara terpisah dan berurutan melalui strategi pipeline untuk memastikan ketepatan setiap tahapan, namun tetap mengandalkan satu model deep learning seperti BERT agar seluruh informasi dapat menyatu dan menghasilkan analisis akurat. Penelitian ini menerapkan strategi *Two-Stage Pipeline Training* menggunakan model IndoBERT, di mana tahap pertama berfokus pada *sequence labeling* dengan skema *BIO Tagging* untuk mengidentifikasi entitas aspek, opini, dan *intensifier*. Tahap kedua melakukan klasifikasi sentimen untuk menentukan polaritas dari pasangan aspek, opini, dan *intensifier*. Semua komponen yang diproses secara terpisah dan kemudian dipelajari secara bersamaan melalui *fine-tuning*, sehingga model mampu memahami konteks hubungan antar entitas dan menentukan polaritas sentimen secara otomatis dan akurat.

2.1.5 Pemodelan

Penelitian ini menggunakan model *indobenchmark/indobert-base-pl* dengan konstruksi input *sentence-pair* [CLS] kalimat [SEP] aspek-opini-intensifier untuk mempelajari hubungan ulasan dan *triplet* sentimen. Proses pelatihan dilakukan dengan pendekatan *fine-tuning*, yaitu menyesuaikan parameter model terhadap dataset penelitian. Model dioptimasi dengan optimizer AdamW dengan learning rate yang diatur ke $3e-5$. Secara spesifik, tingkat dropout pada attention dan hidden layer ditingkatkan menjadi 0.2 untuk regularisasi. Proses pelatihan dijalankan dengan batch size 16 dan direncanakan untuk maksimal 7 epoch. Untuk memastikan model tidak *overfitting* dan mendapatkan hasil baik, diterapkan *Early Stopping* dengan patience 2 (berhenti jika tidak ada perbaikan F1-score setelah 2 epoch).

2.1.6 Evaluasi Model

Tahap evaluasi model dilakukan untuk mengukur efektivitas model dalam menyelesaikan dua tugas utama dalam mengenali entitas yang dilakukan dengan membandingkan prediksi label BIO terhadap label sebenarnya. Sejauh mana entitas yang diprediksi model apakah benar-benar akurat sehingga mampu menemukan kembali semua entitas yang seharusnya ada dalam teks, dan ketepatan dalam menentukan polaritas sentimen. Kinerja model dievaluasi menggunakan matrik *Accuracy*, *Precision*, *Recall*, dan *F1-Score* untuk memberikan penilaian performa yang objektif dan menyeluruh.



3. HASIL DAN PEMBAHASAN

3.1 Dataset

Pengumpulan data dilakukan secara sistematis melalui teknik web scraping pada platform media sosial X (Twitter) dengan menetapkan parameter pencarian berbasis kata kunci spesifik yakni 'Program Makan Bergizi Gratis'. Pemilihan kata kunci bertujuan untuk memastikan hubungan data yang mengandung pendapat masyarakat dan kritik terkait pelaksanaan program makan bergizi gratis dari pemerintah. Jumlah data yang berhasil dikumpulkan mencapai 16.263 cuitan tweet yang dapat dilihat pada Tabel 1.

Tabel 1. Contoh Dataset

Tweet
MBG membuat otak cerdas dan lainnya!! Yuk kita dukung. #MakananBergiziGratis https://t.co/x5lhRjaZi5 Makanan Gratis di Bekasi Belum Merata Bagaimana Sekolah Swasta? https://t.co/o6EazGqq5T #sorotberita #makananbergizigratis #kotabekasi #bmmps

Hasil pengumpulan data per bulan pada Tabel 2, membuktikan bahwa proses web scraping berhasil menerapkan parameter waktu dari Januari hingga Juni 2025. Sistem secara konsisten menangkap distribusi diskusi publik yang relevan dengan kata kunci tersebut di setiap bulannya. Tercatat pada periode Januari hingga Mei, jumlah data yang terkumpul berkisar antara 1.265 hingga 1.095 cuitan, sebelum akhirnya melonjak hingga mencapai 10.076 cuitan pada bulan Juni. Hal ini membuktikan kemampuan sistem dalam melakukan ekstraksi data secara konsisten berdasarkan parameter kata kunci yang telah diterapkan.

Tabel 2. Distribusi Jumlah Tweet per Bulan

Bulan	Jumlah
Januari	1.265
Februari	2.018
Maret	1.029
April	780
Mei	1.095
Juni	10.076

3.2 Pra-pemrosesan Data

Pada tahapan pra-pemrosesan ini berfungsi untuk mengubah data teks asli yang tidak terstruktur menjadi format yang bersih dan bermakna agar dapat diproses secara optimal oleh model. Proses ini bertujuan untuk membersihkan (*cleaning*) teks dari URL, tanda baca, hashtag, dan sebagainya sehingga model hanya berfokus pada kalimat yang mengandung makna. Hasil pra-pemrosesan data bisa dilihat pada Tabel 3.

Tabel 3. Contoh Data Setelah *Cleaning*

Text Asli	Text Bersih
MBG membuat otak cerdas dan lainnya!! Yuk kita dukung. #MakananBergiziGratis https://t.co/x5lhRjaZi5 Makanan Gratis di Bekasi Belum Merata Bagaimana Sekolah Swasta? https://t.co/o6EazGqq5T #sorotberita #makananbergizigratis #kotabekasi #bmmps	makan bergizi gratis membuat otak cerdas dan lainnya yuk kita dukung makananbergizigratis makanan gratis di beksi belum merata bagaimana sekolah swasta sorotberita makananbergizigratis kotabekasi bmmps

Tabel 3 menyajikan perbandingan antara teks asli dengan hasil dari tahapan pra-pemrosesan. Proses ini berhasil menghapus tanda baca, URL, hashtag, dan simbol aneh, serta melakukan *case folding* dengan mengonversi seluruh karakter menjadi huruf kecil untuk menyeragamkan format teks. Selain menghapus teks yang tidak relevan, proses pra-pemrosesan juga melakukan deduplikasi untuk menghapus ulasan yang memiliki isi teks sama persis. Langkah ini bertujuan untuk mencegah model mempelajari pola berulang yang dapat menyebabkan *overfitting*. Keberhasilan proses ini dibuktikan dengan berkurangnya jumlah data tweet dari 16.263 data mentah menjadi 12.400 (23,75%) setelah melakukan tahap pra-pemrosesan.

3.3 Pelabelan BIO Tagging

Meskipun data telah melalui tahap pra-pemrosesan dan tokenisasi, data tersebut masih berupa kumpulan token yang belum memiliki anotasi label. Oleh karena itu, diperlukan proses pelabelan data agar dapat digunakan dalam pembelajaran model untuk mengidentifikasi setiap kelas entitas dalam teks. Pelabelan ini bertujuan untuk menentukan kata atau frasa yang merepresentasikan tiga jenis entitas, yaitu aspek, opini, dan *intensifier* dalam suatu ulasan. Proses ini diawali dengan tahapan *Sequence Labeling* menggunakan skema BIO (Begin, Inside, Outside). Setiap kalimat dipecah menjadi token, kemudian setiap token diberi label BIO dimana label (B) digunakan untuk menandai awal entitas, (I) untuk menandai lanjutan entitas, dan (O) untuk di luar entitas. Dengan pendekatan ini, model tidak hanya melakukan klasifikasi kalimat

secara utuh tetapi mampu mengidentifikasi aspek, opini, dan intensifier secara detail sehingga dapat diketahui apa yang dibahas dan bagaimana opini tersebut ditegaskan dalam sebuah kalimat. Contoh hasil pelabelan *BIO Tagging* dapat dilihat pada Tabel 4.

Tabel 4. Hasil Pelabelan *BIO Tagging*

Kalimat	B-ASP	I-ASP	B-OPI	I-OPI	B-INT
pantas program makan bergizi gencar banget buat diimplementasikan segera ya eh bagaimana kawal mbg	program	Makan bergizi	gratis	gencar	banget

Tabel 4 menunjukkan hasil proses pelabelan data menggunakan skema *BIO Tagging*, yang menjadi tahap awal dari proses pelatihan (*training*) model. Bertujuan untuk menentukan kata dan istilah yang mewakili aspek, opini dan intensifier. Proses pelabelan ini dilakukan secara semi-manual dengan menerapkan berbasis aturan (*Rule-Based*). Melalui penerapan ini, memungkinkan model mengenali pola hubung antar-entitas secara akurat.

Hasil dari proses *sequence labeling* tersebut kemudian menjadi acuan untuk menguji konsistensi model dalam mengenali pola antar-pelabelan di seluruh dataset. Melalui analisis *confusion matrix* terhadap label asli (*ground truth*), dari proses inilah yang membentuk agar model mengenali semua entitas aspek, opini, dan intensifier dalam teks. Untuk melihat sejauh mana performa model dalam mengenali ketiga entitas tersebut, dilakukan pengujian *Confusion Matrix* pada data sampel sebanyak 2.480 ulasan dengan total pemrosesan 66.136 token. Berdasarkan hasil pengujian yang disajikan pada Gambar 2, terlihat bahwa performa model sangat stabil dalam memprediksi ketiga kelas entitas utama melalui pola kesepakatan pada proses pelabelan *BIO tagging*. Meskipun menunjukkan performa sangat baik, masih ditemukan pola ambiguitas berupa kesalahan prediksi pada beberapa entitas tertentu.



Gambar 2. Confusion Matrix *BIO Tagging*

Seperti yang terlihat pada Gambar 2, menunjukkan hasil perbandingan jumlah pola prediksi benar dan kesalahan prediksi untuk setiap entitas. Detail dari pola tersebut dijelaskan sebagai berikut:

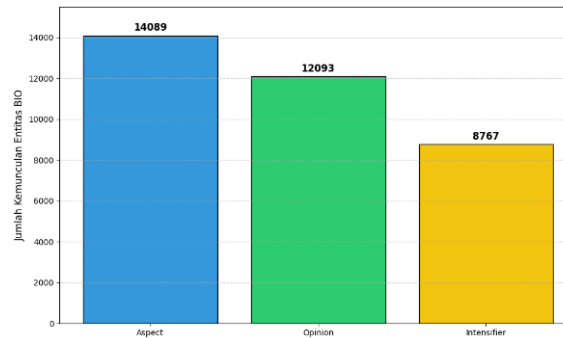
a. Tingkat Kesepakatan (Prediksi Benar)

1. Kelas 'O' (Outside): Sebanyak 21.432 token berhasil diprediksi dengan benar. Ini menunjukkan model sangat stabil dalam mengenali kata-kata yang tidak memiliki muatan entitas.
2. Kelas 'B-ASP' (Begin-Aspect): Sebanyak 14.127 token berhasil diidentifikasi sebagai awal dari sebuah aspek.
3. Kelas 'I-ASP' (Inside-Aspect): Sebanyak 16.486 token diprediksi dengan tepat. Hal ini menandakan sangat fasih dalam mengenali kelanjutan frasa aspek yang panjang.
4. Kelas 'B-OPI' (Begin-Opinion): Sebanyak 6.343 token berhasil diprediksi dengan benar sebagai awal dari kata opini. Kemampuan ini menunjukkan bahwa model sangat efektif dalam membedakan kata sifat yang bermuatan sentimen (bagus, kecewa, setuju) dari kata keterangan.
5. Kelas 'I-OPI' (Inside-Opinion): Sebanyak 553 token diprediksi dengan tepat sebagai kelanjutan dari frasa opini, keberhasilan ini menandakan model mampu menangani opini yang terdiri dari lebih dari satu kata secara konsisten.
6. Kelas 'B-INT' (Begin-Intensifier): Model berhasil mengenali 2.897 token intensitas dengan tepat, yang menjadi kunci keberhasilan dalam klasifikasi kata penegas (banget, amat, sekali).

b. Pola ambiguitas (Kesalahan Prediksi)

1. Kebingungan antara kelas B-ASP dan I-ASP: Tercatat sebanyak 635 kasus dimana label yang seharusnya I-ASP diprediksi sebagai B-ASP. Hal ini menunjukkan adanya tantangan pada aspek yang terdiri dari banyak kata, di mana model menganggap kata di tengah sebagai awal aspek baru.
2. Ambiguitas Entitas terhadap Kelas Luar O: Terdapat 539 kasus di mana entitas aspek dianggap sebagai kata non-entitas (O), dan sebaliknya ada 477 kasus di mana kata non-entitas dianggap B-ASP. Ini mengindikasikan adanya penggunaan istilah kosakata umum yang sering kali muncul dalam konteks aspek.
3. Keterkaitan label B-OPI & I-ASP: Terdapat 267 kasus di mana model keliru memprediksi label B-OPI sebagai I-ASP. Hal ini menunjukkan bahwa dalam ulasan, kata sifat sering kali dianggap menyatu dengan nama aspek itu sendiri.
4. Misklasifikasi pada label B-INT: Terdapat 200 kasus di mana kata non-entitas (O), diklasifikasikan sebagai B-INT. Hal ini kemungkinan disebabkan oleh kata-kata penguat yang bersifat umum namun tidak merujuk pada intensitas opini dalam konteks kalimat tertentu.

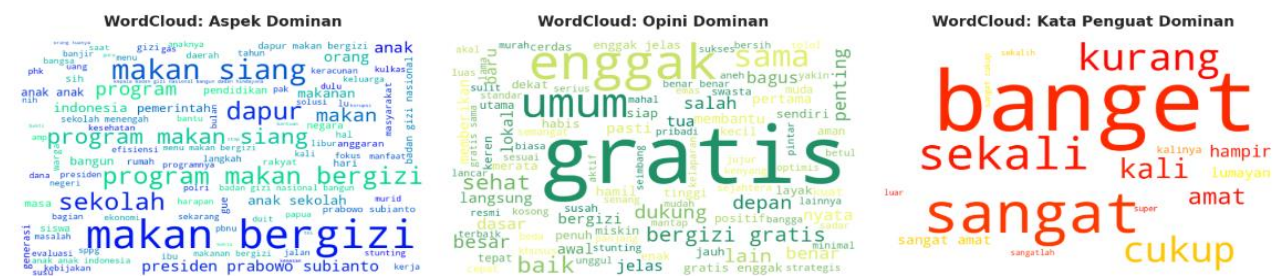
Melalui *Confusion Matrix Ground Truth*, terlihat bahwa model tidak hanya unggul dalam deteksi aspek dan opini, tetapi juga berhasil mengidentifikasi entitas *Intensifier*. Keberhasilan ini menjadi poin krusial dalam penelitian karena label tersebut merupakan indikator utama dalam menegaskan polaritas opini publik. Berdasarkan hasil distribusi kelas kemunculan entitas pada Gambar 3, jumlah token *Intensifier* (8.797) relatif lebih sedikit dibandingkan dengan token aspek (14.089) dan opini (12.093). Kondisi ini terjadi karena data yang digunakan diperoleh melalui proses scraping dari platform Twitter berdasarkan kata kunci tertentu, sehingga tidak semua teks mengandung kata menekankan (*intensifier*). Secara alami, keberadaan entitas *intensifier* dalam kalimat bersifat opsional dan tidak selalu muncul dalam setiap opini yang disampaikan pengguna.



Gambar 3. Distribusi Kelas Entitas

Ketidakeimbangan distribusi antarkelas entitas pada Gambar 3 ini tidak tergolong ekstrem, sehingga masih dapat dikelola. Kondisi ini justru mengonfirmasi perlunya penerapan metode stratifikasi (*stratify*) pada tahap pembagian data latih dan data uji. Stratifikasi bertujuan menjaga proporsi kelas minoritas (dalam hal ini entitas *intensifier*) tetap terwakili secara adil di kedua subset data, sehingga evaluasi performa model dapat dilakukan secara objektif. Selain itu, diterapkan juga kombinasi strategi penanganan ketidakseimbangan kelas, meliputi: *oversampling* terhadap kelas minoritas melalui pendekatan resampling data, *undersampling* terhadap kelas mayoritas, serta *class weighting* dalam fungsi loss untuk memastikan model tidak terpengaruh secara berlebihan oleh kelas dominan.

Untuk memvalidasi kualitas data hasil *BIO Tagging* dengan menggunakan *word cloud* bertujuan untuk memastikan bahwa label yang diberikan sesuai dengan *ground truth* secara semantik. Dari hasil analisis, kata-kata dominan pada setiap kelas entitas (aspek, opini, *intensifier*) adalah kata-kata inti bermakna emosional, bukan kata-kata fungsional. Hal ini membuktikan bahwa proses pelabelan berhasil dan dataset layak digunakan guna melatih model dalam menentukan polaritas sentimen, seperti yang ditunjukkan pada Gambar 4.

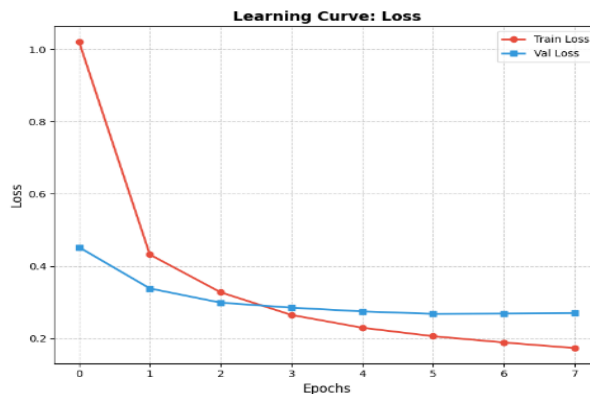


Gambar 4. Visualisasi *WordCloud* Kelas Entitas

Hasil visualisasi pada Gambar 4 memberikan informasi penting. Pertama, keberhasilan model mampu menentukan entitas aspek, opini, dan berhasil mengidentifikasi entitas *Intensifier*. Ini didorong oleh proses pelabelan *BIO Tagging* yang memberikan panduan struktural bagi model untuk mengenali pola kata dan istilah yang mewakili setiap kelas entitas dalam kalimat. Kedua, validitas dataset terbukti, memudahkan analisis data teks (NLP) untuk membedakan ribuan kelas entitas yang relevan untuk menentukan makna mendalam dari setiap ulasan mengenai topik yang ingin di analisis.

3.4 Pemodelan

Proses pelatihan model dievaluasi melalui *Learning Curve* menunjukkan hasil yang cukup baik. Dimana grafik ini memplot nilai kesalahan model pada data latih (*Training Loss*) dan data validasi (*Validation Loss*) di setiap akhir epoch, seperti yang ditunjukkan pada Gambar 5.



Gambar 5. Grafik Training dan Validation Loss

Seperti yang terlihat pada Gambar 5, proses evaluasi performa pemodelan dilakukan melalui analisis kurva *Training* dan *Validation Loss* selama 7 epoch. Dan hasil kedua kurva, baik *Train Loss* (merah) maupun *Val Loss* (biru) sama-sama menunjukkan tren menurun yang konsisten terutama setelah epoch ke-3. Dimana kedua kurva bergerak sangat berdekatan dan tidak menunjukkan adanya celah (*gap*) yang melebar. Hal ini menunjukkan bahwa model tidak mengalami *overfitting*, melainkan memiliki kemampuan generalisasi yang sangat baik dalam membedakan intensitas opini. Proses pelatihan yang stabil ini merupakan hasil dari penerapan hyperparameter yang tepat, termasuk peningkatan dropout rate menjadi 0.2 dan penggunaan learning rate yang lebih rendah (3e-5), yang berhasil mencegah model terlalu menghafal data latih.

3.5 Distribusi Sentimen per Aspek

Dalam analisis sentimen berbasis aspek (ABSA), sentimen tidak dianalisis secara keseluruhan melainkan menganalisis sentimen berdasar topik yang dibahas dari sebuah teks. Pendekatan ini memungkinkan pemahaman yang lebih mendalam mengenai kritik dan respons terhadap suatu topik tertentu, seperti yang ditunjukkan pada Tabel 5. Tabel ini menampilkan sepuluh aspek utama yang paling sering dibahas pada persepsi publik terhadap topik 'program makan bergizi gratis'.

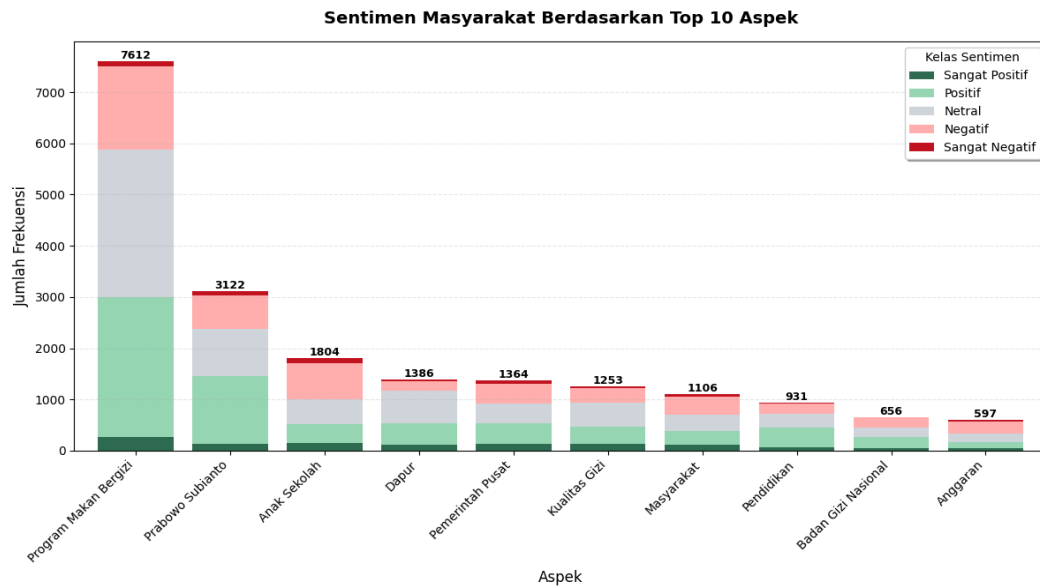
Tabel 5. Hasil Distribusi Sentimen Per Aspek

Aspek	Positif	Netral	Negatif	Sangat Positif	Sangat Negatif	Total
program makan bergizi	2720	2881	1633	274	104	7.612
prabowo subianto	1320	926	658	132	86	3.122
anak sekolah	358	495	702	154	95	1.804
dapur	408	637	189	118	34	1.386
pemerintah pusat	411	388	375	126	64	1.364
kualitas gizi	337	469	295	128	24	1.253
masyarakat	275	323	345	108	55	1.106
pendidikan	381	272	197	63	18	931
badan gizi nasional	215	184	195	52	10	656
anggaran	132	159	231	42	33	597

Berdasarkan hasil dari Tabel 5, penerapan proses pelabelan *BIO tagging* pada IndoBERT berhasil mengidentifikasi kata dan istilah yang mewakili persepsi publik terhadap program makan bergizi gratis dari sepuluh aspek utama. Dimana aspek 'program makan bergizi' menjadi topik paling banyak dibicarakan dengan total (7.612) tweet ini mencerminkan apresiasi publik terhadap konsep program yang akan dijalankan dan diikuti oleh aspek 'prabowo subianto' (3.122) sebagai tokoh figur penggagas kebijakan program. Selanjutnya, aspek 'anak sekolah' mencatat (1.804) tweet yang mengindikasikan perhatian publik terhadap pemerataan manfaat program di lingkungan sekolah. Untuk aspek 'dapur' mencatat (1.386) tweet menyoroti kesiapan infrastruktur teknis dalam mendukung pelaksanaan program. Sementara itu aspek 'pemerintah pusat' (1.364) merefleksikan sorotan terhadap peran dan tanggung jawab pemerintah dalam merealisasikan kebijakan tersebut di setiap daerah.

Selain membahas kebijakan dan menjalankan program, perhatian publik juga menyoroti pembahasan 'kualitas gizi' yang cukup banyak dibahas dengan total (1.253) tweet menjadi perbincangan terkait standar nutrisi dan manfaat kesehatan yang diharapkan dari program tersebut. Adapun aspek 'masyarakat' mencatat (1.106) tweet yang menggambarkan keterlibatan dan harapan publik dalam pelaksanaan program. Berbeda dengan aspek 'pendidikan' dengan (931) tweet yang mengindikasikan apresiasi publik terhadap kontribusi program dalam mencerdaskan anak

bangsa melalui pemenuhan gizi di lingkungan sekolah. Terakhir, aspek ‘badan gizi nasional’ sebanyak (656) tweet yang menyoroti kinerja lembaga baru sebagai pengelola program. Sedangkan aspek ‘anggaran’ dengan (597) tweet terkait dengan pembiayaan program dan transparansi pengelolaan dana. Dari setiap aspek yang telah teridentifikasi, model tidak hanya berfungsi untuk mengenali setiap entitas aspek dalam sebuah kalimat tetapi juga menganalisis sentimen terhadap topik yang dibicarakan dalam teks. Secara keseluruhan dari hasil distribusi sentimen per aspek, publik lebih mendukung konsep kebijakan ‘Makan Bergizi Gratis’ dibandingkan dengan berbagai kelemahan dalam pelaksanaannya. Serta berhasil mengidentifikasi intensitas opini (*Intensifier*) yang mencerminkan dukungan sangat kuat maupun kritik tajam terhadap kinerja lembaga baru tersebut. Untuk memvalidasi kualitas dataset yang telah diproses dapat dilihat dari visualisasi frekuensi distribusi sentimen per aspek pada Gambar 6.



Gambar 6. Distribusi Sentimen Per Aspek

Hasil distribusi pada Gambar 6 menunjukkan keberhasilan proses *BIO tagging* dalam mengidentifikasi entitas aspek, opini, maupun menemukan entitas baru *intensifier*. Berbeda dengan analisis konvensional, penggunaan label *intensifier* ini memungkinkan model mampu menangkap kedalaman emosi masyarakat, mulai dari kritik tajam (Sangat Negatif) hingga apresiasi tinggi (Sangat Positif) dalam memetakan berbagai aspek yang menjadi perhatian publik. Meskipun pada gradasi warna yang lebih rendah dibandingkan kategori lainnya, membuktikan model masih mampu menyaring adanya kata menekankan dalam opini masyarakat.

3.4 Evaluasi Model

Pada implementasi skema pelabelan *BIO Tagging* dalam mengidentifikasi kata atau frasa yang mewakili aspek, opini, serta *intensifier* menunjukkan hasil yang sangat baik dalam mengenali konteks ulasan. Berdasarkan hasil pengujian yang disajikan pada Tabel 6, model yang telah melalui proses *fine-tuning* secara keseluruhan berhasil mencapai tingkat akurasi sebesar 94% dengan rata-rata *F1-Score* mencapai 0.94. Pencapaian ini mengindikasikan bahwa model memiliki performa yang sangat baik dan konsisten dalam mengidentifikasi seluruh kelas entitas secara presisi dan seimbang dalam mengenali seluruh kelas entitas.

Tabel 6. Hasil Evaluasi BIO Tagging

BIO tagging	Precision	Recall	F1-score
B-ASP	0.94	0.92	0.95
B-INT	0.92	0.93	0.93
B-OPI	0.94	0.95	0.95
I-ASP	0.92	0.95	0.92
I-OPI	0.92	0.91	0.93
O	0.91	0.92	0.92
accuracy			0.94
macro avg	0.93	0.92	0.92
weighted avg	0.94	0.94	0.94

Hasil evaluasi model pada Tabel 6 menunjukkan stabilitas performa yang cukup baik di seluruh kelas entitas pada pelabelan *BIO tagging*, di mana nilai *F1-score* untuk setiap kelas berkisar antara (0.91) hingga (0.95). Hal ini membuktikan bahwa proses *fine-tuning* tidak mengalami penurunan performa dan bekerja dengan sangat konsisten dalam mengenali struktur kalimat untuk menemukan kata pada entitas aspek, opini dan *intensifier* (intensitas opini). Dalam



identifikasi entitas aspek, pada label ‘B-ASP’ menunjukkan performa sangat baik dengan *F1-Score* (0.95) maupun label ‘I-ASP’ sebesar (0.92) yang mengidentifikasi ketepatan model dalam menemukan kata awal dan kata lanjutan yang merepresentasikan entitas aspek dalam sebuah kalimat. Sementara itu, pada entitas opini, menunjukkan nilai *F1-Score* yang sangat tinggi pada awal kata yang menyampaikan penilaian terhadap aspek ‘B-OPI’ sebesar 0.95 serta performa yang tetap stabil pada kata lanjutan ‘I-OPI’ dengan skor (0.93).

Poin yang paling krusial, yaitu pada entitas *intensifier* ‘B-INT’ sebagai penanda kata penegasan terhadap penilaian terhadap opini, menunjukkan hasil yang cukup baik dengan *F1-Score* sebesar (0.93). Hal ini menunjukkan bahwa model mampu menemukan kata yang memperkuat maupun penekanan terhadap kalimat opini. Selain itu, pada label ‘O’ (Outside) untuk menandai kata yang bukan merupakan bagian dari entitas menunjukkan hasil *F1-score* cukup tinggi berkisar antara (0.92). Hal ini membuktikan bahwa model memiliki kemampuan yang sangat baik dalam membedakan antara informasi yang relevan bagi sistem dengan kata-kata umum lainnya dalam teks.

Setelah ketiga entitas berhasil diidentifikasi melalui proses *BIO tagging*, keberhasilan ekstraksi entitas ini menjadi fondasi utama dalam menentukan polaritas sentimen berdasarkan hubungan antar-entitas. Untuk mengetahui seberapa baik model dalam memutuskan apakah suatu pasangan (aspek, opini, dan *intensifier*) saling terkait, hasil evaluasinya disajikan pada Tabel 7. Tabel ini menyajikan evaluasi kemampuan model dalam mengklasifikasikan polaritas sentimen. Berbeda dengan analisis sentimen konvensional yang hanya membedakan sentimen positif dan negatif, penelitian ini menambahkan entitas *intensifier* sehingga klasifikasi diperluas menjadi lima kategori: Sangat Positif, Positif, Netral, Negatif, dan Sangat Negatif. Dengan adanya *intensifier*, model tidak hanya mengenali pendapat yang disampaikan, tetapi juga mengukur tingkat intensitas emosional dalam kalimat.

Tabel 7. Hasil Evaluasi Sentimen

Sentimen	Precision	Recall	F1-score
Sangat Negatif	0.82	0.84	0.82
Negatif	0.85	0.84	0.87
Netral	0.86	0.82	0.89
Positif	0.87	0.94	0.92
Sangat Positif	0.84	0.87	0.85
accuracy			0.88
macro avg	0.85	0.85	0.85
weighted avg	0.88	0.88	0.88

Berdasarkan hasil evaluasi model pada Tabel 7, model berhasil mencapai akurasi sebesar 88% dengan nilai *Weighted Average F1-Score* mencapai 0.88. Hal ini menunjukkan bahwa model memiliki performa yang sangat baik dalam menentukan hubungan yang saling berkaitan antara entitas aspek, opini, dan *intensifier* secara akurat dalam menentukan polaritas sentimen. Dalam evaluasi sentimen pada tabel tersebut, menunjukkan efektivitas cukup baik dalam mengklasifikasikan kelima kelas sentimen. Ditandai dengan *F1-score* kelas ‘Positif’ yang memperoleh nilai tertinggi (0.92), diikuti oleh kelas ‘Netral’ (89), ‘Negatif’ (87). Pada klasifikasi sentimen yang melibatkan hubungan antar-entitas aspek dan intensitas opini (*intensifier*), model tetap mempertahankan akurasi yang sangat baik dengan *F1-score* (0.85) kelas ‘Sangat Positif’ serta nilai (0.82) untuk kelas ‘Sangat Negatif’. Secara keseluruhan, hasil evaluasi menunjukkan bahwa model terbukti efektif dalam mengklasifikasikan polaritas sentimen ke dalam lima kategori. Termasuk membedakan tingkat intensitas opini (*intensifier*) baik untuk memperkuat atau memperlemah dan memberikan nuansa tertentu.

4. KESIMPULAN

Penelitian ini berhasil mengembangkan dan memperluas task ASOTE menjadi empat komponen dengan menyertakan entitas baru yaitu *intensifier* sebagai penanda intensitas opini. Selain itu, penelitian ini mengembangkan sebuah kerangka kerja berbasis *Transformer* dengan penerapan skema *BIO Tagging* yang diperluas dengan penambahan label (B-INT) untuk menandai kata penegas untuk menangani proses ekstraksi pada ASOTE. Berdasarkan hasil evaluasi pada data uji menunjukkan model bekerja sangat baik dalam mengenali kata yang mewakili semua kelas entitas. Sementara itu, pada tahap klasifikasi polaritas sentimen cukup baik dalam menentukan pasangan antar-entitas dengan akurasi sebesar 88% dan *F1-Score* mencapai 0.88. Secara keseluruhan, penelitian ini berhasil mengintegrasikan entitas *intensifier* ke dalam struktur *triplet* ASOTE melalui strategi *Sentence-Pair Classification* mampu mendorong model melakukan penalaran relasional, di mana model dapat menentukan pasangan aspek, opini, dan *intensifier* secara bersamaan untuk menentukan polaritas dari setiap pasangan entitas yang teridentifikasi. Namun, ada beberapa keterbatasan menjadi catatan untuk penelitian berikutnya. Pertama, adanya ambiguitas konteks di mana kata umum (non-entitas) masih terdeteksi sebagai *intensifier* karena maknanya yang terlalu general dalam kalimat tertentu. Kedua, model masih menghadapi tantangan pada *boundary detection*, di mana frasa aspek yang panjang sering kali terpecah dan dianggap sebagai awal entitas baru. Keterbatasan tersebut membuka peluang bagi penelitian selanjutnya untuk menerapkan arsitektur yang lebih kuat dalam mendeteksi batas entitas, seperti penggunaan *Conditional Random Fields* (CRF). Selain itu, penggunaan *dependency parsing* sangat disarankan agar model dapat lebih akurat dalam membedakan kata penegas yang secara sintaksis merujuk pada opini kebijakan dengan kata keterangan umum.



REFERENCES

- [1] H. Zhao, M. Yang, X. Bai, and H. Liu, "A Survey on Multimodal Aspect-Based Sentiment Analysis," *IEEE Access*, vol. 12, pp. 12039–12052, 2024, doi: 10.1109/ACCESS.2024.3354844.
- [2] K. Scaria, H. Gupta, S. Goyal, S. A. Sawant, S. Mishra, and C. Baral, "InstructABSA: Instruction Learning for Aspect Based Sentiment Analysis," *Proc. 2024 Conf. North Am. Chapter Assoc. Comput. Linguist. Hum. Lang. Technol. NAACL 2024*, vol. 2, pp. 720–736, 2024, doi: 10.18653/v1/2024.naacl-short.63.
- [3] J. G. Carrasco Ramírez, "Natural Language Processing Advancements: Breaking Barriers in Human-Computer Interaction," *J. Artif. Intell. Gen. Sci. ISSN3006-4023*, vol. 3, no. 1, pp. 31–39, 2024, doi: 10.60087/jaigs.v3i1.63.
- [4] M. G. Monika, S. Burra, V. Divya, and S. A. kumar, "Aspect-based Sentimental Analysis for Movie Recommendation," *Int. Sci. J. Eng. Manag.*, vol. 04, no. 01, pp. 1–6, 2025, doi: 10.55041/isjem02242.
- [5] A. P. Al AUFar and A. Romadhony, "Aspect-based Sentiment Analysis on Beauty Product Reviews using BERT and Long Short-Term Memory," *J. Nas. Pendidik. Tek. Inform.*, vol. 14, no. 2, pp. 364–373, 2025, doi: 10.23887/janapati.v14i2.94392.
- [6] Arif Fitra Setyawan, Amelia Devi Putri Ariyanto, Fari Katul Fikriah, and Rozaq Isnaini Nugraha, "Analisis Sentimen Ulasan iPhone di Amazon Menggunakan Model Deep Learning BERT Berbasis Transformer," *Elkom J. Elektron. dan Komput.*, vol. 17, no. 2, pp. 447–452, 2024, doi: 10.51903/elkom.v17i2.2150.
- [7] B. Xu and G. Yang, "Interpretability research of deep learning: A literature survey," *Inf. Fusion*, vol. 115, 2025, doi: 10.1016/j.inffus.2024.102721.
- [8] 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.
- [9] Q. Xu, Y. Hong, J. Chen, J. Yao, and G. Zhou, "Self-augmented sequentiality-aware encoding for aspect term extraction," *Inf. Process. Manag.*, vol. 61, no. 3, 2024, doi: 10.1016/j.ipm.2024.103656.
- [10] H. Wang, C. Wang, C. Li, and C. Wu, "Incorporating Global Information for Aspect Category Sentiment Analysis," *Electron.*, vol. 13, no. 24, 2024, doi: 10.3390/electronics13245020.
- [11] Y. Wen, E. Yu, J. Qu, L. Cheng, Y. Chen, and S. Lu, "Target-Oriented Opinion Words Extraction Based on Dependency Tree," *Big Data Cogn. Comput.*, vol. 9, no. 8, 2025, doi: 10.3390/bdcc9080207.
- [12] G. Liu, Y. Zhao, and B. Shen, "A Span-based Enhanced Bidirectional Extraction Framework for Multi-word Aspect Sentiment Triplets," *J. Internet Technol.*, vol. 26, no. 2, pp. 199–209, 2025, doi: 10.70003/160792642025032602005.
- [13] X. Li, K. Zhang, and D. Han, "Duality-Driven Aspect Sentiment Triplet Extraction with LLM and Iterative Reinforcement," *Symmetry (Basel)*, vol. 17, no. 5, 2025, doi: 10.3390/sym17050642.
- [14] I. Naglik and M. Lango, "ASTE-Transformer: Modelling Dependencies in Aspect-Sentiment Triplet Extraction," *EMNLP 2024 - 2024 Conf. Empir. Methods Nat. Lang. Process. Find. EMNLP 2024*, pp. 2324–2339, 2024, doi: 10.18653/v1/2024.findings-emnlp.129.
- [15] W. Wu, D. Wang, M. Wang, S. Feng, and Y. Zhang, "Sentiment Triplet Extraction With Multi-View Contrastive Learning," *IEEE Trans. Affect. Comput.*, vol. 16, no. 3, pp. 1526–1542, 2025, doi: 10.1109/TAFFC.2024.3521608.
- [16] J. Peng and B. Su, "Aspect Sentiment Triplet Extraction Based on Deep Relationship Enhancement Networks," *Appl. Sci.*, vol. 14, no. 5, 2024, doi: 10.3390/app14052221.
- [17] R. Mukherjee, N. Kannan, S. Kumar Pandey, and P. Goyal, "CONTRASTE: Supervised Contrastive Pre-training With Aspect-based Prompts For Aspect Sentiment Triplet Extraction," *Find. Assoc. Comput. Linguist. EMNLP 2023*, pp. 12065–12080, 2023, doi: 10.18653/v1/2023.findings-emnlp.807.
- [18] X. Sun, Z. Zhu, J. Qi, Z. Zhao, and H. Pei, "Affective Commonsense Knowledge Enhanced Dependency Graph for aspect sentiment triplet extraction," *J. Supercomput.*, vol. 80, no. 7, pp. 8614–8636, 2024, doi: 10.1007/s11227-023-05778-z.
- [19] J. Liu *et al.*, "Exploiting Duality in Aspect Sentiment Triplet Extraction With Sequential Prompting," *IEEE Trans. Knowl. Data Eng.*, vol. 36, no. 11, pp. 6111–6123, 2024, doi: 10.1109/TKDE.2024.3391381.
- [20] Y. Li, F. Wang, and S. H. Zhong, "A More Fine-Grained Aspect-Sentiment-Opinion Triplet Extraction Task," *Mathematics*, vol. 11, no. 14, 2023, doi: 10.3390/math11143165.
- [21] A. S. Paramita and J. Jusak, "Fine-Grained Sentiment Analysis Approach on Customer Reviews Based on Aspect-Level Emotion Detection," *J. Appl. Data Sci.*, vol. 6, no. 3, pp. 2235–2247, 2025, doi: 10.47738/jads.v6i3.964.