BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 6, No 5, August 2026 | Hal 2100-2108
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v6i5.1321

Analisis Sentimen Berbasis Aspek pada Opini Publik Pascabencana Hidrometeorologi Menggunakan Model Pre-Trained IndoBERT

Muhammad Yusran*, Saidi Ramadan Siregar, Arini Vika Sari

Fakultas Ilmu Komputer dan Teknologi Informasi, Program Studi Teknik Informatika, Universitas Budi Darma, Medan, Indonesia
Email: 1,*myusran804@gmail.com, 2saidiramadan89@gmail.com, 3arinivika1@gmail.com
Email Penulis Korespondensi: myusran804@gmail.com

Abstrak−Provinsi Aceh sangat rentan terhadap bencana, khususnya kejadian hidrometeorologis seperti banjir dan tanah longsor. Data Badan Penanggulangan Bencana Aceh (BPBA) menunjukkan terjadi 418 bencana pada 2023 dengan total kerugian mencapai Rp430 miliar; demikian pula pada 2024, sebagian besar kejadian merupakan bencana hidrometeorologis yang berdampak pada puluhan ribu keluarga. Tingginya frekuensi bencana ini memicu lonjakan reaksi di media sosial, menghasilkan volume data yang tidak dapat dianalisis secara mendalam menggunakan metode analisis sentimen konvensional. Penelitian ini menggunakan metode Aspect-Based Sentiment Analysis (ABSA) dan model pre-trained IndoBERT (indobenchmark/indobert-base-p1) untuk menganalisis opini publik terkait lima aspek penanggulangan bencana: logistik dan bantuan, evakuasi dan tempat penampungan, koordinasi pemerintah, infrastruktur dan rekonstruksi, serta layanan kesehatan. Sebanyak 1.000 opini dikumpulkan melalui X (Twitter) dan YouTube (November 2025 – April 2026); data diproses melalui pra-pemrosesan teks menggunakan pustaka indoNLP, menghasilkan 3.319 pasangan opini-aspek berlabel. Model dikembangkan dengan klasifikasi pasangan kalimat dan di-fine-tune menggunakan CrossEntropy Loss (learning rate 2e-5, batch size 16, 3 epoch), dengan evaluasi kinerja berdasarkan akurasi, presisi, recall, dan skor F1 (nilai target >85%). Hasil pengujian menunjukkan model IndoBERT mencapai akurasi 91,26% dan skor F1 sebesar 86,03%, melampaui target kinerja. Analisis distribusi sentimen mengungkapkan aspek layanan kesehatan memicu sentimen paling positif (89,82%), mencerminkan apresiasi publik terhadap upaya tim medis. Sebaliknya, sentimen negatif mendominasi aspek evakuasi dan tempat penampungan (34,45%), mengindikasikan ketidakpuasan masyarakat terhadap prosedur evakuasi dan ketersediaan tempat penampungan pascabencana di Aceh.

Kata Kunci: Aspect-Based Sentiment Analysis; IndoBERT; Opini Publik; Bencana Hidrometeorologi; Aceh

Abstract−Aceh Province is highly vulnerable to disasters, particularly hydrometeorological events like floods and landslides. Aceh Disaster Management Agency (BPBA) data shows 418 disasters occurred in 2023 with total losses reaching Rp430 billion; similarly, in 2024, most incidents were hydrometeorological disasters affecting tens of thousands of families. This high frequency triggered a social media reaction surge, producing data volumes unanalyzable via conventional sentiment methods. This study applies Aspect-Based Sentiment Analysis (ABSA) and the pre-trained IndoBERT model (indobenchmark/indobert-base-p1) to analyze public opinion across five disaster management aspects: logistics and assistance, evacuation and shelters, government coordination, infrastructure and reconstruction, and healthcare. A total of 1,000 opinions were gathered from X (Twitter) and YouTube (November 2025–April 2026), pre-processed via indoNLP, yielding 3,319 labeled opinion-aspect pairs. The model used sentence-pair classification, fine-tuned with CrossEntropy Loss (learning rate 2e-5, batch size 16, 3 epochs), evaluated via accuracy, precision, recall, and F1-score (target >85%). Testing shows IndoBERT achieved 91.26% accuracy and an 86.03% F1-score, exceeding targets. Sentiment distribution reveals healthcare triggered the most positive sentiment (89.82%), reflecting public appreciation for medical teams. Conversely, negative sentiment dominated evacuation and shelters (34.45%), indicating public dissatisfaction with evacuation procedures and shelter availability post-disaster in Aceh.

Keywords: Aspect-Based Sentiment Analysis; IndoBERT; Public Opinion; Hydrometeorological Disasters; Aceh

1. PENDAHULUAN

Sebagai negara kepulauan yang terletak di garis khatulistiwa dengan topografi yang kompleks, Indonesia menghadapi tingkat kerawanan yang sangat tinggi terhadap berbagai ancaman bencana alam, khususnya bencana hidrometeorologi seperti banjir, tanah longsor, angin kencang, serta gelombang pasang. Dari sekian banyak wilayah di tanah air, Provinsi Aceh merupakan salah satu daerah yang paling rentan terhadap peristiwa destruktif semacam itu. Akibat letak geografisnya yang strategis namun dihadapkan pada karakteristik iklim serta topografi yang dinamis, wilayah ini sering kali tidak luput dari terjangan bencana berskala besar yang mengancam keselamatan jiwa dan harta benda penduduk.

Data empiris dari Badan Penanggulangan Bencana Aceh (BPBA) menunjukkan tingkat kerentanan yang nyata. Sepanjang tahun 2023, tercatat tidak kurang dari 418 kejadian bencana di provinsi tersebut yang memicu total kerugian ekonomi hingga mencapai Rp430 miliar, di mana mayoritas kerugian didominasi oleh dampak destruktif dari bencana banjir dan tanah longsor. Situasi darurat ini semakin diperparah pada tahun berikutnya. Berdasarkan data BPBA tahun 2024, puluhan ribu keluarga yang mencakup hingga 44.641 rumah tangga teridentifikasi menjadi korban terdampak bencana alam, yang sebagian besar kembali dikategorikan sebagai bencana hidrometeorologi.

Dampak kemanusiaan ini terus berlanjut tanpa henti. Hanya dalam beberapa bulan pertama tahun 2025 (periode Januari hingga Mei), tercatat 138 kejadian bencana baru yang memilukan, bahkan merenggut 272 nyawa penduduk setempat. Tingginya frekuensi dan eskalasi dampak bencana yang berulang ini tidak hanya menjadi krisis kemanusiaan di lapangan, tetapi juga memicu gelombang reaksi, emosi, serta kritik keras dari masyarakat luas yang disuarakan secara masif melalui berbagai platform media sosial.

Di era digital saat ini, masyarakat semakin aktif memanfaatkan ruang-ruang digital untuk membagikan respons, keluh kesah, kritik, maupun apresiasi terhadap berbagai peristiwa penting, termasuk penanganan krisis akibat bencana alam. Platform-platform media sosial ini menyimpan kekayaan informasi berupa opini publik yang sangat berharga [1].

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 6, No 5, August 2026 | Hal 2100-2108
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v6i5.1321

Data teks tersebut memuat sentimen nyata mengenai berbagai dimensi respons pascabencana mulai dari efektivitas logistik dan distribusi bantuan, prosedur evakuasi, koordinasi antarlembaga pemerintah, hingga pemulihan infrastruktur publik yang rusak [1][2]. Kendati demikian, pengolahan data opini publik dalam jumlah besar (big data) menghadirkan tantangan tersendiri. Metode analisis sentimen konvensional selama ini memiliki keterbatasan mendasar, yaitu hanya mampu mengukur polaritas sentimen secara global atau totalitas (seperti mengklasifikasikan teks ke dalam kategori positif, negatif, atau netral) tanpa mampu membedah sentimen terhadap aspek-aspek spesifik yang dibahas di dalam teks tersebut [3].

Sebagai solusi atas kendala tersebut, *Aspect-Based Sentiment Analysis* (ABSA) hadir sebagai pendekatan canggih dalam ranah *Natural Language Processing* (NLP) [4]. ABSA mampu mengidentifikasi dan mengekstraksi sentimen yang bernuansa secara spesifik terhadap target aspek tertentu dalam suatu kalimat [5][6]. Metode ini dinilai jauh lebih komparatif dan informatif karena mampu menyajikan pemetaan masalah yang terperinci. Dengan ABSA, pemangku kebijakan dapat mengetahui secara presisi aspek mana dari penanganan pascabencana yang mendapat respons positif ataupun menuai kritik negatif dari publik. Perkembangan teknologi kecerdasan buatan, khususnya arsitektur berbasis Transformer seperti BERT *(Bidirectional Encoder Representations from Transformers)*, telah merevolusi cara mesin memahami bahasa manusia dan memberikan terobosan masif dalam bidang NLP. Untuk konteks bahasa Indonesia, IndoBERT muncul sebagai model BERT pra-latih *(pre-trained)* yang dirancang khusus dan dilatih menggunakan korpus teks bahasa Indonesia berskala raksasa [7].

Dibandingkan dengan model multibahasa umum seperti mBERT, IndoBERT terbukti memiliki keunggulan komparatif yang signifikan dalam hal pemahaman konteks semantik, struktur tata bahasa, serta penanganan ragam bahasa informal atau slang yang kerap dijumpai dalam teks percakapan di media sosial [10][11]. Hal ini menjadikan IndoBERT sebagai fondasi model yang paling ideal dan relevan untuk menganalisis opini publik berbahasa Indonesia secara akurat [9].

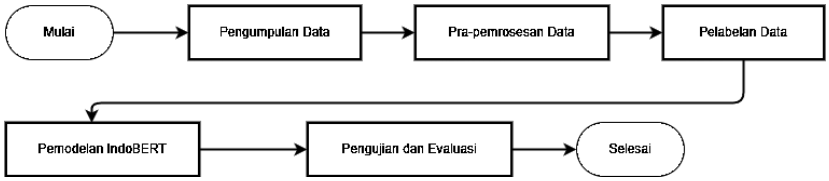
Beberapa penelitian sebelumnya telah mengkaji penerapan ABSA di berbagai sektor. Devlin dkk. memperkenalkan arsitektur BERT, yang menjadi tonggak penting bagi model *pre-trained* modern [10]. Selanjutnya, Wilie dkk. mengembangkan model IndoBERT dan tolok ukur IndoNLU, yang menunjukkan bahwa model monolingual untuk bahasa Indonesia memiliki kinerja lebih baik dibandingkan varian multilingual [11]. Terkait penerapannya, Yulianti dan Nissa menerapkan ABSA menggunakan IndoBERT pada ulasan pelanggan; mereka menemukan bahwa model yang telah melalui proses *fine-tuning* untuk klasifikasi pasangan kalimat mengungguli arsitektur *deep learning* terdahulu seperti *Word2Vec* dengan pencapaian skor F1 yang hingga 23,6% lebih tinggi [12]. Pendekatan *fine-tuning* IndoBERT secara *end-to-end* juga terbukti meningkatkan kinerja klasifikasi aspek dibandingkan metode hibrida konvensional [13], dan penggunaan BERT dengan konstruksi kalimat bantu *(auxiliary sentence)* juga menjadi terobosan yang efektif. Purwitasari dkk. menyusun dataset berbasis aspek dari media sosial mengenai kebijakan pemerintah [14], sementara Prasetyo dkk. menganalisis data Twitter terkait bencana alam di Indonesia [15]; namun, penelitian mereka terbatas pada analisis sentimen secara umum dan tidak membahas aspek spesifik penanganan pascabencana.

Tinjauan pustaka ini mengungkap adanya celah penelitian yang mendasar: belum ada penelitian yang secara komprehensif menerapkan ABSA menggunakan model *pre-trained* IndoBERT untuk menganalisis opini publik mengenai penanganan pascabencana hidrometeorologi di Provinsi Aceh

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Tahapan penelitian disusun secara sistematis untuk memastikan pelaksanaan penelitian berjalan sesuai dengan tujuan yang telah ditetapkan. Setiap tahapan saling berkaitan dan dilaksanakan secara berurutan, mulai dari identifikasi permasalahan, studi literatur, pengumpulan data, pengolahan dan analisis data, hingga penyusunan hasil penelitian. Alur tahapan tersebut menjadi dasar pelaksanaan penelitian secara terarah dan sistematis. Secara visual, tahapan penelitian disajikan pada Gambar 1.



Gambar 1. Tahapan Penelitian

Berikut adalah penjelasan rinci mengenai setiap tahapan kerangka penelitian pada Gambar 1:

a. Pengumpulan Data *(Data Crawling)*: Yaitu mengekstraksi data opini publik dari platform *X (Twitter)* dan *Youtube* menggunakan API resmi. Penarikan data dari *X (Twitter)* dilakukan menggunakan *Twitter Authentication Token* dan pustaka *tweet-harvest*, dengan parameter kueri pencarian berbasis kata kunci seperti "(#banjiraceh) since:2025-11-15 until:2026-04-30". Sementara itu, pengumpulan data dari YouTube memanfaatkan YouTube Data API v3 melalui pustaka googleapiclient.discovery dengan ekstraksi atribut nama pengarang, teks komentar, jumlah suka, dan tanggal publikasi. Seluruh data mentah dari kedua platform digabungkan menggunakan pustaka Pandas, lalu disaring secara

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 6, No 5, August 2026 | Hal 2100-2108
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v6i5.1321

manual untuk menyingkirkan duplikat, spam, retweet, serta teks mengenai bencana non hidrometeorologi (seperti gempa bumi). Proses penyaringan ini menghasilkan 1.000 opini publik yang representatif dalam rentang waktu November 2025 hingga April 2026.

b. Pra-Pemrosesan Data *(Text Processing)*: Menggunakan pustaka indoNLP meliputi *Case Folding (lowercase)*, *Data Cleansing* (menghapus URL, *mention*, tagar, karakter non-alfabet menggunakan RegEx), dan *Word Normalization* (mengubah slang menjadi kata baku KBBI).
c. Pelabelan Data *(Data Annotation)*: Mencakup label aspek (lima aspek utama) dan label sentimen (positif, netral, negatif).
d. Pemodelan Menggunakan IndoBERT: dengan pendekatan *sentence-pair classification*. Total 3.319 pasangan data yang dihasilkan dibagi menggunakan teknik *Stratified Holdout* dengan rasio 80:20, di mana 80% data (2.655 pasangan) digunakan untuk pelatihan *(training)* dan 20% data (664 pasangan) untuk pengujian *(testing)*. Dari data pelatihan, 10% dialokasikan kembali sebagai data validasi selama proses *fine-tuning*.
e. Pengujian dan Evaluasi: Hasil dievaluasi menggunakan Confusion Matrix dan metrik akurasi, presisi, recall, serta F1-score.

2.2 Formulasi Matematis Pemodelan IndoBERT

Pemodelan analisis sentimen berbasis aspek dalam penelitian ini mengadopsi *Sentence-Pair Classification* berbasis model pre-trained IndoBERT. Input terdiri dari teks ulasan (w) dan teks aspek (a) yang ditokenisasi menggunakan *WordPiece Tokenizer* menjadi sekuens X dengan penambahan token spesial [CLS] dan [SEP] [12].

$$X = ([CLS], w_1, \dots, w_n, [SEP], a_1, \dots, a_m, [SEP]) \quad (1)$$

Persamaan (1) menunjukkan bahwa teks ulasan dan teks aspek digabungkan menjadi satu sekuens input yang terdiri atas token [CLS], token ulasan, token [SEP], token aspek, dan token [SEP]. Selanjutnya, sekuens tersebut diteruskan ke *Transformer Encoder* untuk memperoleh representasi kontekstual. Pada tahap ini, model menghitung keterkaitan antar-token menggunakan mekanisme *Multi-Head Self-Attention* dengan matriks *Query* (Q), *Key* (K), dan *Value* (V) [16].

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (2)$$

Persamaan (2) menunjukkan bahwa mekanisme *Multi-Head Self-Attention* digunakan untuk menghitung bobot keterkaitan antar-token berdasarkan matriks *Query* (Q), *Key* (K), dan *Value* (V), sehingga menghasilkan representasi kontekstual dari sekuens input. Model mengekstrak vektor *hidden state* dari token [CLS] yang merangkum keseluruhan konteks, lalu diteruskan ke lapisan linear *(Dense layer)* dengan fungsi aktivasi Softmax untuk menghitung probabilitas sentimen [10].

$$P(y \mid X) = softmax\left(W \cdot h_{[CLS]} + b\right) \quad (3)$$

Persamaan (3) menunjukkan bahwa vektor *hidden state* dari token [CLS] digunakan sebagai representasi keseluruhan konteks untuk menghasilkan probabilitas pada setiap kelas sentimen melalui lapisan linear dan fungsi aktivasi *Softmax*. Model dilatih ulang *(fine-tuning)* untuk memperbarui bobot parameter. Minimalisasi kesalahan prediksi dihitung menggunakan fungsi *Cross-Entropy Loss* [17].

$$L = -\sum_{c=1}^{C} y_c \log(\hat{y}_c) \quad (4)$$

Persamaan (4) menunjukkan bahwa fungsi *Cross-Entropy Loss* digunakan untuk mengukur kesalahan antara label sentimen sebenarnya dan probabilitas sentimen yang diprediksi oleh model. Nilai *loss* tersebut menjadi dasar dalam proses pembaruan parameter model selama *fine-tuning* [18].

2.2 Arsitektur dan Implementasi Sistem

Sistem ABSA dalam penelitian ini dirancang sebagai alur kerja modular yang mengintegrasikan berbagai komponen pemrosesan. Komponen-komponen tersebut meliputi tahap pra-pemrosesan data menggunakan pustaka indoNLP, modul deteksi aspek berbasis kata kunci, tokenisasi melalui *BertTokenizer*, serta model IndoBERT yang dikombinasikan dengan lapisan klasifikasi. Arsitektur sistem menggunakan kerangka kerja Django dengan pola MVT *(Model-View-Template)*, serta didukung oleh Tailwind CSS dan Chart.js untuk visualisasi data.

Dari segi fungsionalitas, sistem ini mencakup enam kasus penggunaan utama: mengakses halaman utama, memasukkan teks berisi opini, mengajukan permintaan analisis, melihat hasil analisis sentimen berbasis aspek, meninjau grafik probabilitas dan ringkasan hasil, serta memeriksa status aplikasi. Proses prediksi dimulai ketika pengguna memasukkan teks berisi opini melalui antarmuka. Sistem memvalidasi input tersebut dan kemudian mengajukan permintaan analisis. Selanjutnya, sistem melakukan deteksi kata kunci untuk mengidentifikasi aspek-aspek yang ada secara iteratif. Teks dan nama aspek diproses oleh model IndoBERT untuk memprediksi probabilitas sentimen dan mengklasifikasikannya sebagai negatif, netral, atau positif.

Implementasi sistem dibagi menjadi dua tahap utama yaitu pelatihan model *(fine-tuning)* dan pengembangan aplikasi. Pelatihan model dilakukan pada platform Google Colab menggunakan akselerasi GPU T4. Spesifikasi perangkat

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 6, No 5, August 2026 | Hal 2100-2108
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v6i5.1321

keras mencakup prosesor Intel Xeon (2 vCPU dengan kecepatan 2,2 GHz), RAM 16 GB, GPU T4, dan penyimpanan 100 GB. Spesifikasi perangkat lunak mencakup sistem operasi Linux Ubuntu 22.04, bahasa pemrograman Python 3.11, kerangka kerja pembelajaran mesin PyTorch 2.1+ dan Hugging Face Transformers 4.36+, serta format penyimpanan model Safetensors.

3. HASIL DAN PEMBAHASAN

3.1 Pra-Pemrosesan dan Deteksi Aspek

Untuk pra-pemrosesan data teks dalam penelitian ini, digunakan fungsi-fungsi bawaan dari pustaka indoNLP—yang dikembangkan secara khusus untuk bahasa Indonesia. Proses ini sangat penting karena data teks dari platform media sosial seperti X (Twitter) dan YouTube sering kali tidak terstruktur, mengandung banyak derau *(noise)*, serta memuat berbagai karakter yang tidak relevan. Prosedur ini dimulai dengan tahap *Case Folding*, yaitu mengubah semua karakter alfabet menjadi huruf kecil untuk menyeragamkan format. Selanjutnya, data dibersihkan menggunakan modul *Regular Expression* (RegEx) guna menghapus elemen yang tidak memiliki nilai semantik, seperti URL, *mention* (@), tagar (#), angka, dan tanda baca. Pola RegEx seperti re.sub(r'[^a-z\s]', ' ', text diterapkan untuk memastikan teks hanya berisi karakter alfabet dan spasi, sementara kelebihan spasi, tab, dan pemisah baris dihapus menggunakan re.sub(r'\s+', ' ', text).strip(). Tahap akhir adalah normalisasi kata, di mana teks dipecah menjadi token kata yang kemudian dibandingkan dengan kamus normalisasi. Kamus ini memuat singkatan dan istilah non-baku (bahasa gaul) yang umum ditemukan di media sosial seperti mengubah "yg" menjadi "yang", "gk" menjadi "tidak", dan "udh" menjadi "sudah" sehingga memastikan teks sepenuhnya mematuhi kaidah standar bahasa Indonesia sebagaimana ditetapkan dalam KBBI.

Fungsi bawaan dari pustaka indoNLP digunakan pada tahap pra-pemrosesan data teks. Berdasarkan hasil pra-pemrosesan, diketahui bahwa seluruh 1.000 data opini bebas dari *noise*. Setiap aspek dalam penelitian ini didefinisikan oleh sekumpulan kata kunci yang relevan; secara keseluruhan, terdapat 117 kata kunci yang tersebar di lima aspek berbeda. Hasil deteksi menunjukkan bahwa 699 dari 1.000 opini (69,9%) memuat setidaknya satu aspek, berdasarkan pencocokan kata kunci. Rincian mengenai definisi aspek dan daftar kata kunci yang terkait disajikan dalam Tabel 1.

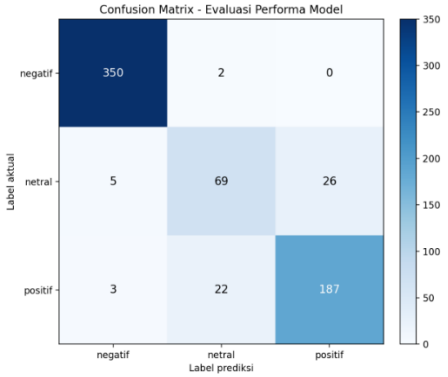
Tabel 1. Kata Kunci Per Aspek

Aspek	Jumlah Kata Kunci	Contoh Kata Kunci
Logistik dan Bantuan	23	sembako, makanan, pakaian, tenda, donasi, bantuan, logistik
Evakuasi dan Pengungsian	20	evakuasi, sar, basarnas, pengungsi, relokasi, shelter
Koordinasi Pemerintah	28	pemerintah, bpbd, bnpb, gubernur, bupati, kebijakan
Infrastruktur dan Rekonstruksi	24	jalan, jembatan, listrik, rekonstruksi, longsor
Layanan Kesehatan	22	kesehatan, dokter, obat, puskesmas, ambulans

Berdasarkan Tabel 1, aspek 'Koordinasi Pemerintah' memiliki jumlah kata kunci terbanyak (28), sedangkan aspek 'Evakuasi dan Pengungsian' memiliki jumlah paling sedikit (20). Setiap opini yang diberi label secara manual dikonversi menjadi *tuple* data dengan format (opini, aspek, label). Melalui proses penataan ulang ini, diperoleh total 3.319 pasangan opini-aspek yang kemudian disiapkan untuk digunakan. Proses pelatihan model dikonfigurasi menggunakan modul `TrainingArguments` dari pustaka *Hugging Face*, dengan menerapkan hiperparameter seperti model dasar `indobenchmark/indobert-base-p1`, arsitektur `BertForSequenceClassification`, panjang urutan maksimum 128 token, ukuran *batch* 16, laju pembelajaran 2e-5, pengoptimal AdamW, dan 3 *epoch*.

3.2 Pengujian dan Evaluasi Model

Evaluasi kinerja model IndoBERT dilakukan menggunakan set pengujian yang mencakup 20% dari total pasangan opini-aspek, yang dipilih melalui teknik *Stratified Holdout*, sehingga menghasilkan 664 sampel uji. Hasilnya dianalisis menggunakan matriks konfusi yang menampilkan jumlah prediksi benar dan salah untuk setiap kelas sentimen (negatif, netral, dan positif). Visualisasi nilai-nilai dari matriks konfusi tersebut disajikan pada Gambar 2.



Gambar 2. *Confusion Matrix* Evaluasi Performa Model

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 6, No 5, August 2026 | Hal 2100-2108
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v6i5.1321

Berdasarkan Gambar 2, nilai *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN) dapat ditentukan untuk setiap kelas sentimen. Untuk kelas negatif, model memprediksi dengan benar 350 titik data (TP), sementara 2 titik data netral diprediksi secara keliru sebagai negatif (FP) dan tidak ada titik data positif yang dilabeli secara keliru sebagai negatif. Pada kelas netral, 69 titik data diprediksi dengan benar, namun 5 titik data negatif dan 26 titik data positif diklasifikasikan secara keliru sebagai netral. Untuk kelas positif, model memprediksi dengan benar 187 titik data, dengan 3 titik data negatif dan 22 titik data netral diprediksi secara keliru sebagai positif. Secara keseluruhan, model memprediksi dengan benar 606 dari 664 titik data uji dan melakukan 58 kesalahan prediksi.

Berdasarkan matriks konfusi pada Gambar 2, metrik evaluasi untuk setiap kelas dihitung secara rinci. Mengingat penelitian ini melibatkan klasifikasi multikelas dengan ketidakseimbangan kelas, pendekatan *Macro-Averaging* [19] digunakan untuk menghitung presisi, *recall*, dan skor F1.

a. Untuk Kelas Negatif

$$TP = 350$$

$$FP = 5 + 3 = 8$$

$$FN = 2 + 0 = 2$$

$$Presisi = \frac{350}{350+8} = 97{,}77\%$$

$$Recall = \frac{350}{350+2} = 99{,}43\%$$

$$F1 - Score = \frac{2\times(97{,}77\%\times99{,}43\%)}{(97{,}77\%+99{,}43\%)} = 98{,}59\%$$

b. Untuk Kelas Netral

$$TP = 69$$

$$FP = 2 + 22 = 24$$

$$FN = 5 + 26 = 31$$

$$Presisi = \frac{69}{69+24} = 74{,}19\%$$

$$Recall = \frac{69}{69+31} = 69{,}00\%$$

$$1 - Score = \frac{2\times(74{,}19\%\times69{,}00\%)}{(74{,}19\%+69{,}00\%)} = 71{,}50\%$$

c. Untuk Kelas Positif

$$TP = 187$$

$$FP = 0 + 26 = 26$$

$$FN = 3 + 22 = 25$$

$$Presisi = \frac{187}{187+26} = 87{,}79\%$$

$$Recall = \frac{187}{187+25} = 88{,}21\%$$

$$F1 - Score = \frac{2\times(87{,}79\%\times88{,}21\%)}{(87{,}79\%+88{,}21\%)} = 88{,}00\%$$

d. Rata-rata *(Macro)* dan Akurasi Keseluruhan

$$Macro\ Presisi = \frac{97{,}77\%+74{,}19\%+87{,}79\%}{3} = 86{,}58\%$$

$$Macro\ Recall = \frac{99{,}43\%+69{,}00\%+88{,}21\%}{3} = 85{,}55\%$$

$$Macro\ F1 - Score = \frac{(98{,}59\%+71{,}50\%+88{,}00\%)}{3} = 86{,}03\%$$

$$Akurasi\ Keseluruhan = \frac{350+69+187}{64} = \frac{606}{664} = 91{,}26\%$$

Berdasarkan hasil perhitungan metrik evaluasi pada masing-masing kelas, diperoleh nilai presisi, *recall*, dan F1-*Score* untuk kelas negatif, netral, dan positif. Nilai dari ketiga metrik tersebut kemudian dihitung menggunakan

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 6, No 5, August 2026 | Hal 2100-2108
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v6i5.1321

pendekatan *Macro-Averaging* untuk memperoleh nilai rata-rata kinerja model pada seluruh kelas. Sementara itu, akurasi digunakan untuk mengukur tingkat ketepatan prediksi model secara keseluruhan. Hasil perhitungan metrik evaluasi tersebut disajikan pada Tabel 2.

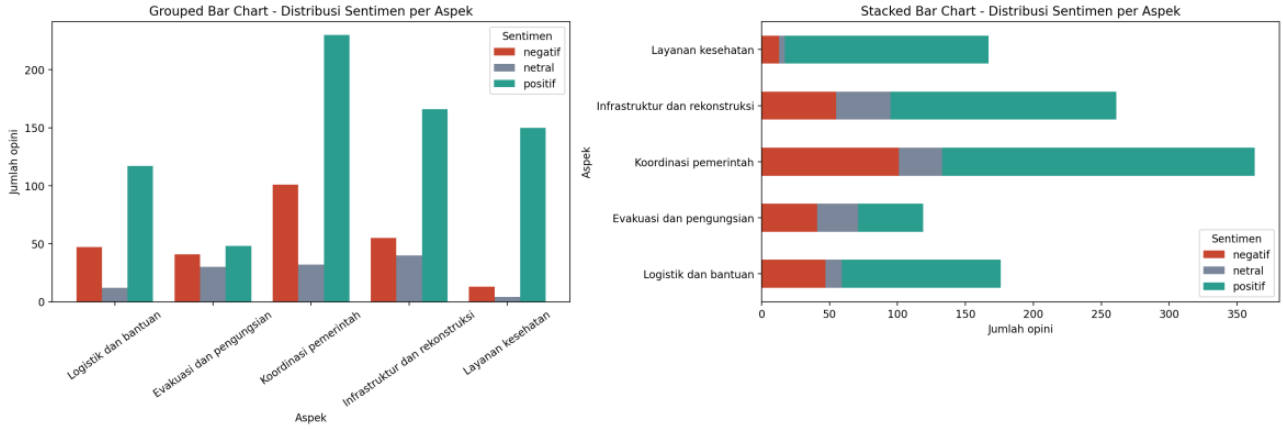
Tabel 2. Hasil Matriks Evaluasi Model

Matriks	Nilai (%)
Akurasi	91,26
Presisi (Macro Avg)	86,58
Recall (Macro Avg)	85,55
F1-*Score* (Macro Avg)	86,03

Berdasarkan Tabel 2, model IndoBERT memperoleh akurasi sebesar 91,26%, yang menunjukkan bahwa sebagian besar data berhasil diklasifikasikan dengan benar. Nilai presisi Macro Avg sebesar 86,58% menunjukkan bahwa model memiliki kemampuan yang baik dalam menghasilkan prediksi yang sesuai pada ketiga kelas sentimen. Sementara itu, nilai recall Macro Avg sebesar 85,55% menunjukkan bahwa model mampu mengenali sebagian besar data pada masing-masing kelas sentimen. Nilai F1-Score Macro Avg sebesar 86,03% menunjukkan keseimbangan yang baik antara presisi dan recall pada seluruh kelas. Secara keseluruhan, nilai F1-Score sebesar 86,03% telah melampaui target kinerja yang ditetapkan, yaitu di atas 85%. Hasil tersebut menunjukkan bahwa model IndoBERT memiliki kinerja klasifikasi yang baik dalam mengidentifikasi sentimen pada pasangan aspek-opini dalam konteks penanggulangan bencana.

3.3 Analisis Distribusi Sentimen per Aspek

Setelah reliabilitas model dipastikan, model tersebut digunakan untuk memprediksi sentimen pada keseluruhan dataset. Sebanyak 1.086 pasangan aspek-opini berhasil diidentifikasi dari 3.319 pasangan yang terdapat dalam data pelatihan. Untuk memberikan gambaran visual yang lebih komprehensif mengenai distribusi data, distribusi opini per aspek dipetakan menggunakan diagram batang berkelompok dan bertumpuk, sebagaimana ditunjukkan pada Gambar 3.



Gambar 3. Visualisasi Distribusi Sentimen per Aspek

Diagram pada Gambar 3 menunjukkan distribusi jumlah opini berdasarkan kelas sentimen negatif, netral, dan positif pada lima aspek manajemen bencana. Secara umum, sentimen positif mendominasi seluruh aspek, dengan jumlah tertinggi terdapat pada aspek koordinasi pemerintah, diikuti oleh infrastruktur dan rekonstruksi serta layanan kesehatan. Sementara itu, aspek logistik dan bantuan juga menunjukkan dominasi sentimen positif meskipun jumlah opininya lebih rendah. Pada aspek evakuasi dan pengungsian, jumlah opini positif relatif lebih sedikit dibandingkan aspek lainnya, tetapi tetap menjadi sentimen yang paling dominan. Sentimen negatif paling banyak ditemukan pada aspek koordinasi pemerintah dan infrastruktur dan rekonstruksi, sedangkan sentimen netral cenderung memiliki jumlah opini yang lebih rendah pada seluruh aspek. Pola tersebut menunjukkan bahwa persepsi positif masyarakat relatif lebih dominan dalam pembahasan berbagai aspek manajemen bencana. Rincian mengenai distribusi persentase sentimen berdasarkan aspek disajikan dalam Tabel 3.

Tabel 3. Distribusi Sentimen per Aspek

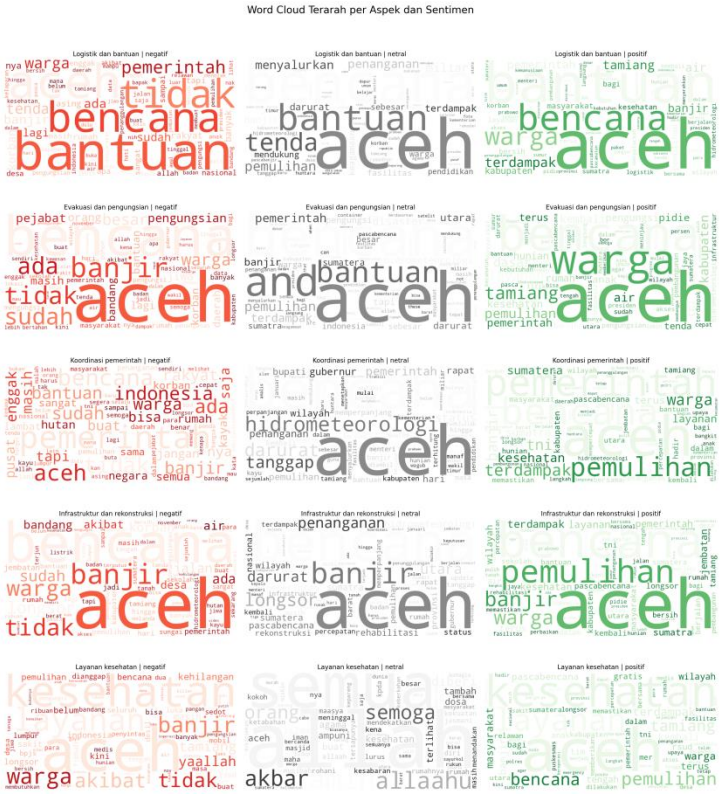
Aspek	Jumlah Data	Positif (%)	Netral (%)	Negatif (%)
Logistik dan Bantuan	176	67,05	6,82	26,14
Evakuasi dan Pengungsian	119	37,82	27,73	34,45
Koordinasi Pemerintah	363	62,53	9,64	27,82
Infrastruktur dan Rekonstruksi	261	63,60	15,33	21,07
Layanan Kesehatan	167	89,82	2,40	7,78

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 6, No 5, August 2026 | Hal 2100-2108
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v6i5.1321

Berdasarkan Tabel 3, sentimen positif mendominasi pada seluruh aspek manajemen bencana, dengan proporsi tertinggi pada aspek layanan kesehatan sebesar 89,82%. Sebaliknya, aspek evakuasi dan pengungsian memiliki proporsi sentimen positif terendah, yaitu 37,82%, serta proporsi sentimen negatif tertinggi sebesar 34,45%. Kondisi tersebut menunjukkan bahwa respons masyarakat terhadap aspek evakuasi dan pengungsian cenderung lebih beragam dibandingkan aspek lainnya. Pada aspek logistik dan bantuan, sentimen positif mencapai 67,05%, sedangkan sentimen netral dan negatif masing-masing sebesar 6,82% dan 26,14%. Aspek koordinasi pemerintah menunjukkan 62,53% sentimen positif, 9,64% netral, dan 27,82% negatif. Sementara itu, infrastruktur dan rekonstruksi memiliki 63,60% sentimen positif, 15,33% netral, dan 21,07% negatif. Adapun layanan kesehatan menunjukkan distribusi sentimen paling positif dengan 89,82% sentimen positif, 2,40% netral, dan hanya 7,78% negatif. Secara keseluruhan, Tabel 3 menunjukkan bahwa layanan kesehatan merupakan aspek dengan respons positif tertinggi, sedangkan evakuasi dan pengungsian memiliki respons yang relatif lebih beragam serta proporsi sentimen negatif tertinggi. Hasil ini menunjukkan bahwa pendekatan *Aspect-Based Sentiment Analysis* (ABSA) dapat memberikan informasi yang lebih spesifik mengenai kecenderungan persepsi masyarakat pada setiap aspek manajemen bencana dibandingkan analisis sentimen secara keseluruhan.

3.4 Visualisasi Data Opini Publik (*Word Cloud* Terarah)

Untuk melengkapi analisis kuantitatif, dilakukan visualisasi data teks menggunakan *word cloud* yang dikategorikan berdasarkan aspek dan sentimen, sebagaimana ditunjukkan pada Gambar 4. Berbeda dengan *word cloud* standar, visualisasi ini terdiri dari 15 panel yang menampilkan kombinasi antara lima aspek manajemen bencana dan tiga kelas sentimen (negatif, netral, dan positif). Setiap panel pada Gambar 4 merepresentasikan kata-kata yang paling sering muncul pada kombinasi aspek dan kelas sentimen tertentu. Ukuran kata dalam *word cloud* menunjukkan frekuensi kemunculannya, sehingga kata dengan ukuran lebih besar memiliki frekuensi yang lebih tinggi dalam data. Visualisasi ini digunakan untuk mengidentifikasi kata atau istilah yang dominan pada masing-masing aspek dan sentimen. Dengan demikian, *word cloud* terarah dapat memberikan gambaran mengenai karakteristik opini yang muncul pada setiap aspek manajemen bencana berdasarkan kecenderungan sentimennya.



Gambar 4. *Word Cloud* Terarah per Aspek dan Sentimen

Dalam visualisasi tersebut, ukuran teks mencerminkan frekuensi kata, sementara skema warna membedakan kategori sentimen (merah untuk negatif, abu-abu untuk netral, dan hijau untuk positif). Gambar 4 menunjukkan bahwa kata-kata yang paling dominan secara keseluruhan adalah “bencana”, “Aceh”, “banjir”, “bantuan”, “pemerintah", dan “korban”. Tingginya frekuensi kata “banjir” dan “bencana” mencerminkan peristiwa hidrometeorologis yang paling sering dibahas oleh masyarakat di media sosial. Selain itu, membedah visualisasi berdasarkan aspek dan sentimen menghasilkan wawasan kualitatif yang mendalam. Sebagai contoh, pada panel untuk aspek Layanan Kesehatan yang dikaitkan dengan sentimen positif, kata-kata seperti "dokter," "obat-obatan," puskesmas”, dan “pemulihan” muncul secara menonjol. Hal ini mengonfirmasi temuan kuantitatif sebelumnya mengenai tingginya apresiasi masyarakat terhadap

kehadiran tim medis dan ketersediaan obat-obatan. Sebaliknya, panel untuk aspek Evakuasi dan Pengungsian yang dikaitkan dengan sentimen negatif didominasi oleh kata-kata seperti “terlambat”, " pengungsi”, "relokasi," dan "darurat." Visualisasi ini memperkuat bukti adanya ketidakpuasan masyarakat terhadap prosedur evakuasi dan ketersediaan lokasi penampungan. Temuan dari visualisasi *word cloud* ini secara visual mengonfirmasi pemilihan kelima aspek manajemen bencana tersebut serta menunjukkan bahwa pendekatan ABSA berhasil mengekstrak penilaian yang spesifik dan bernuansa dari opini publik di media sosial.

3.5 Pembahasan

Temuan mengenai sentimen negatif yang kuat terkait evakuasi dan fasilitas penampungan sejalan dengan penelitian Imran dkk. [20]. Mereka menunjukkan bahwa koordinasi evakuasi pascabencana sering kali menjadi hambatan kritis dalam upaya bantuan awal akibat kerusakan infrastruktur dan terputusnya akses jalan. Dalam konteks Aceh, medan yang sulit dan tingginya frekuensi banjir bandang menyebabkan keterlambatan tambahan dalam pemindahan korban, yang memicu reaksi negatif di media sosial. Prasetyo dkk. [15] juga menekankan bahwa keterlambatan evakuasi secara konsisten menimbulkan kekecewaan publik, sebagaimana dibuktikan oleh data Twitter.

Sebaliknya, sentimen yang sangat positif terkait layanan kesehatan (89,82%) menunjukkan efektivitas upaya bantuan medis pascabencana. Hal ini mengindikasikan bahwa kebijakan pemerintah mengenai pengerahan tim medis dan distribusi obat-obatan selaras dengan pilar-pilar rencana penanggulangan bencana nasional [2]. Perbandingan temuan-temuan ini menegaskan pentingnya penelitian ini, karena metode ABSA memungkinkan pembedaan antara masalah operasional di lapangan (seperti lambatnya evakuasi) dan hasil positif (seperti layanan kesehatan). Analisis sentimen konvensional pada tingkat dokumen tidak dapat menangkap nuansa semacam itu [21]; oleh karena itu, temuan ini memberikan rekomendasi yang lebih terarah bagi BPBA untuk memprioritaskan evaluasi prosedur pencarian dan penyelamatan (SAR) serta ketersediaan tempat penampungan darurat.

4. KESIMPULAN

Berdasarkan rumusan masalah, tujuan penelitian, serta hasil analisis dan pengujian yang dilakukan dalam studi ini, dapat disimpulkan bahwa identifikasi aspek manajemen terkait pemulihan pascabencana hidrometeorologi di Aceh berhasil dicapai menggunakan metode deteksi berbasis kata kunci. Dari 1.000 opini publik yang dikumpulkan, 699 di antaranya secara eksplisit memuat setidaknya satu dari lima aspek manajemen, sehingga menghasilkan total 3.319 kombinasi opini-aspek yang valid. Model IndoBERT (indobenchmark/indobert-base-p1) berhasil diimplementasikan dan dievaluasi menggunakan metode klasifikasi pasangan kalimat; model ini terbukti efektif dalam memahami konteks bahasa Indonesia informal di media sosial, dengan mencapai akurasi sebesar 91,26% dan F1-score sebesar 86,03%. Sentimen publik terhadap berbagai aspek respons pascabencana berhasil dianalisis; hasilnya menunjukkan bahwa aspek evakuasi dan pengungsian memiliki proporsi sentimen negatif tertinggi (34,45%) yang mengindikasikan ketidakpuasan publik sementara aspek layanan kesehatan menunjukkan tingkat sentimen positif tertinggi (89,82%), yang mencerminkan apresiasi kuat dari masyarakat. Keterbatasan penelitian ini meliputi fakta bahwa pelabelan data dilakukan oleh satu orang anotator saja dan deteksi aspek hanya mengandalkan pencocokan kata kunci; oleh karena itu, penelitian selanjutnya sebaiknya melibatkan lebih dari satu anotator untuk memungkinkan penghitungan reliabilitas antar-penilai *(inter-rater reliability)* serta memanfaatkan metode ekstraksi aspek otomatis, seperti *dependency parsing* atau *Named Entity Recognition* (NER).

REFERENCES

[1] B. Ilyas and A. Sharifi, “A systematic review of social media-based sentiment analysis in disaster risk management,” *Int. J. Disaster Risk Reduct.*, vol. 123, p. 105487, Jun. 2025, doi: 10.1016/J.IJDRR.2025.105487.

[2] Rencana Nasional Penanggulangan Bencana 2020-2024, “Rencana Nasional Penanggulangan Bencana 2020-2024,” Jakarta, 2021. Accessed: May 04, 2026. [Online]. Available: https://www.bnpb.go.id/buku/rencana-nasional-penanggulangan-bencana-20202024

[3] E. Cambria, Q. Liu, S. Decherchi, F. Xing, and K. Kwok, “SenticNet 7: A Commonsense-based Neurosymbolic AI Framework for Explainable Sentiment Analysis,” Marseille: European Language Resources Association, 2022, pp. 3829–3839. Accessed: May 04, 2026. [Online]. Available: https://aclanthology.org/2022.lrec-1.408/

[4] D. Khurana, A. Koli, K. Khatter, and S. Singh, “Natural Language Processing: State of The Art, Current Trends and Challenges,” *Multimed. Tools Appl.*, vol. 82, no. 3, pp. 3713–3744, Jan. 2023, doi: 10.1007/s11042-022-13428-4.

[5] G. Brauwers and F. Frasincar, “A Survey on Aspect-Based Sentiment Classification,” *ACM Comput. Surv.*, vol. 55, no. 4, Mar. 2022, doi: 10.1145/3503044.

[6] W. Zhang, X. Li, Y. Deng, L. Bing, and W. Lam, “Towards Generative Aspect-Based Sentiment Analysis,” in *ACL-IJCNLP 2021 - 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, Proceedings of the Conference*, Association for Computational Linguistics (ACL), 2021, pp. 504–510. doi: 10.18653/v1/2021.acl-short.64.

[7] N. M. Gardazi, A. Daud, M. K. Malik, A. Bukhari, T. Alsahfi, and B. Alshemaimri, “BERT Applications in Natural Language Processing: A Review,” *Artif. Intell. Rev.*, vol. 58, no. 166, pp. 1–49, Mar. 2025, doi: 10.1007/s10462-025-11162-5.

[8] B. Wilie *et al.*, “IndoNLU: Benchmark and Resources for Evaluating Indonesian Natural Language Understanding,” in *Proceedings of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing, AACL-IJCNLP 2020*, K. F. Wong, K. Knight, and H. Wu, Eds.,

Suzhou, China: Association for Computational Linguistics (ACL), 2020, pp. 843–857. doi: 10.18653/v1/2020.aacl-main.85.

[9] M. F. Mubaraq and W. Maharani, "Sentiment Analysis on Twitter Social Media Towards Climate Change on Indonesia Using IndoBERT Model," *J. MEDIA Inform. BUDIDARMA*, vol. 6, no. 4, p. 2426, Oct. 2022, doi: 10.30865/mib.v6i4.4570.

[10] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," in *NAACL HLT 2019 - 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies - Proceedings of the Conference*, J. Burstein, C. Doran, and T. Solorio, Eds., Stroudsburg, PA, USA: Association for Computational Linguistics, 2019, pp. 4171–4186. doi: 10.18653/V1/N19-1423.

[11] N. Istiqomah and F. Novika, "Comparative Performance of IndoBERT and IndoLEM Baseline Models for Post-Disaster Health Information Extraction from Indonesian Online News," *J. Comput. Sci. Informatics Eng.*, vol. 4, no. 3, pp. 158–174, Jul. 2025, doi: 10.55537/COSIE.V4I3.1174.

[12] 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. 3579–3589, Oct. 2024, doi: 10.11591/eei.v13i5.8032.

[13] T. Aprilah, D. Rosal, I. M. Setiadi, and W. Herowati, "Enhancing Aspect-Based Sentiment Analysis via Hugging Face Fine-Tuned IndoBERT," *J. Appl. Informatics Comput.*, vol. 9, no. 6, pp. 3821–3830, Dec. 2025, doi: 10.30871/JAIC.V9I6.11409.

[14] D. Purwitasari, C. B. P. Putra, and A. B. Raharjo, "A Stance Dataset with Aspect-Based Sentiment Information from Indonesian COVID-19 Vaccination-Related Tweets," *Data Br.*, vol. 47, pp. 1–16, Apr. 2023, doi: 10.1016/j.dib.2023.108951.

[15] V. R. Prasetyo, G. Erlangga, and D. A. Prima, "Analisis Sentimen untuk Identifikasi Bantuan Korban Bencana Alam berdasarkan Data di Twitter Menggunakan Metode K-Means dan Naive Bayes," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 10, no. 5, pp. 1055–1062, Oct. 2023, doi: 10.25126/jtiik.2023107077.

[16] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, "Attention Is All You Need," in *Attention Is All You Need*, Long Beach, CA, USA: Curran Associates, Inc., 2017, pp. 5998–6008. doi: 10.48550/arXiv.1706.03762.

[17] Suhariyanto, R. Sarno, C. Fatichah, and R. Abdullah, "Aspect-based Sentiment Analysis: Natural Language Understanding for Implicit Review," *Int. J. Electr. Comput. Eng.*, vol. 14, no. 6, pp. 6711–6722, 2024, doi: 10.11591/ijece.v14i6.pp6711-6722.

[18] I. Duru and A. S. Sunar, "Transformer and Pre-Transformer Model-Based Sentiment Prediction with Various Embeddings : A Case Study on Amazon Reviews," *Entropy*, vol. 27, no. 12, p. 1202, 2025, doi: https://doi.org/10.3390/e27121202.

[19] M. Grandini, E. Bagli, and G. Visani, "Metrics for Multi-Class Classification: an Overview," Aug. 2020, Accessed: Aug. 07, 2026. [Online]. Available: https://arxiv.org/pdf/2008.05756

[20] M. Imran, F. Ofli, D. Caragea, and A. Torralba, "Using AI and Social Media Multimodal Content for Disaster Response and Management: Opportunities, Challenges, and Future Directions," *Inf. Process. Manag.*, vol. 57, no. 5, p. 102261, Sep. 2020, doi: 10.1016/j.ipm.2020.102261.

[21] M. Wankhade, A. C. S. Rao, and C. Kulkarni, "A survey on sentiment analysis methods, applications, and challenges," *Artif. Intell. Rev. 2022 557*, vol. 55, no. 7, pp. 5731–5780, Feb. 2022, doi: 10.1007/S10462-022-10144-1.