
ASPECT BASED SENTIMENT ANALYSIS TERHADAP PERTAMINA PADA PLATFORM X MENGGUNAKAN INDOBERTWEET

Muhamad Fazri¹⁾, Nono Heryana²⁾, Ahmad Khusaeri³⁾

Program Studi Sistem Informasi, Fakultas Ilmu Komputer

Universitas Singaperbangsa Karawang

Jl. H.S. Ronggowaluyo, Telukjambe Timur, Karawang, Jawa Barat 41361

e-mail: muhamadfazri758@gmail.com¹⁾, nono@unsika.ac.id²⁾, ahmad.khusaeri@cs.unsika.ac.id³⁾

Abstrak

Penelitian ini bertujuan menganalisis opini publik terhadap Pertamina melalui platform X menggunakan pendekatan Aspect-Based Sentiment Analysis (ABSA) dengan model IndoBERTweet. Data cuitan berbahasa Indonesia dikumpulkan periode September–Oktober 2025 menggunakan tweet-harvest v2.6.1 dengan kata kunci "Pertamina lang:id". Penelitian menggunakan kerangka Knowledge Discovery in Database (KDD) yang mencakup pengumpulan data, preprocessing, ekstraksi aspek, klasifikasi sentimen, dan analisis hasil. Ekstraksi aspek menggunakan pendekatan keyword-based menghasilkan lima aspek: Harga BBM, Kualitas BBM, Pelayanan SPBU, Kebijakan Pertamina, dan Umum. Dari 7.486 cuitan mentah diperoleh 6.642 cuitan bersih, kemudian dihasilkan 8.198 data setelah replikasi multi-aspek. Klasifikasi sentimen dilakukan menggunakan model IndoBERTweet pretrained secara otomatis. Hasil menunjukkan sentimen Negatif mendominasi 45,22% (3.707 data), Netral 42,17% (3.457 data), dan Positif 12,61% (1.034 data). Aspek Kebijakan Pertamina (55,4%) dan Kualitas BBM (55,0%) mencatat proporsi negatif tertinggi, sedangkan aspek Umum memiliki distribusi paling berimbang dengan positif tertinggi (28,2%). Temuan ini digunakan sebagai dasar rekomendasi strategis peningkatan layanan dan komunikasi publik Pertamina.

Kata kunci : Analisis Sentimen Berbasis Aspek; IndoBERTweet; Media Sosial; Pertamina; Platform X

1. Pendahuluan

Perkembangan teknologi informasi dan media sosial telah mengubah cara masyarakat menyampaikan opini secara publik, terutama melalui platform mikroblogging seperti X (sebelumnya Twitter) yang memungkinkan pengguna berbagi komentar secara *real-time* [1]. Indonesia pada tahun 2023 tercatat memiliki lebih dari 167 juta pengguna aktif media sosial (60,4% dari total populasi), menjadikan media sosial kanal utama diskusi publik [2]. Pertamina sebagai BUMN pengelola distribusi bahan bakar minyak (BBM) menghadapi tekanan opini publik yang tinggi terkait harga BBM, kualitas produk, pelayanan SPBU, dan kebijakan perusahaan.

Metode analisis sentimen telah diidentifikasi sebagai alat efektif untuk menelaah data teks dari media sosial [3]. Pendekatan *Aspect-Based Sentiment Analysis* (ABSA) memungkinkan pemisahan opini berdasarkan aspek tertentu sehingga wawasan yang diperoleh menjadi lebih spesifik dan *actionable* [4]. Penelitian oleh Said dan Manik (2023) menunjukkan bahwa pemilahan aspek pada data Twitter Indonesia mampu meningkatkan pemahaman terhadap persepsi masyarakat secara signifikan [5].

Perkembangan *Natural Language Processing* (NLP) berbasis transformer, khususnya IndoBERTweet yang dilatih khusus pada teks media sosial berbahasa Indonesia, membuka peluang baru dalam analisis sentimen kontekstual [6]. Namun kajian ABSA berbasis transformer untuk sektor energi nasional Indonesia masih terbatas; sebagian besar penelitian berfokus pada e-commerce dan politik [7]. Penelitian ini mengisi kesenjangan tersebut dengan menerapkan ABSA berbasis IndoBERTweet terhadap opini publik Pertamina di platform X, menghasilkan distribusi sentimen per aspek sebagai dasar rekomendasi strategis.

2. Landasan Teori

Analisis Sentimen

Analisis sentimen merupakan pendekatan komputasional untuk mengidentifikasi dan mempelajari opini subjektif dari teks [8]. Dalam konteks media sosial, pendekatan ini mengonversi data teks tidak terstruktur menjadi wawasan kuantitatif tentang persepsi dan sikap pengguna terhadap suatu entitas. Terdapat tiga level granularitas: dokumen, kalimat, dan aspek. Penelitian ini beroperasi di level aspek yang memberikan pemahaman paling granular [9]. Dominasi sentimen negatif pada platform seperti X umumnya disebabkan *negativity bias* di mana pengguna lebih terdorong membagikan pengalaman buruk dibandingkan pengalaman positif, terutama untuk layanan publik yang berdampak langsung pada kehidupan sehari-hari.

Aspect-Based Sentiment Analysis (ABSA)

ABSA merupakan ekstensi dari analisis sentimen umum yang mengidentifikasi aspek dari suatu entitas beserta sentimen yang terkait [10]. Dua sub-tugas utama adalah *Aspect Term Extraction* (ATE) untuk mengidentifikasi frasa aspek, dan *Aspect Sentiment Classification* (ASC) untuk menentukan polaritas sentimen per aspek. Penelitian ini mengadopsi pendekatan *keyword-based* untuk ATE dan model pretrained IndoBERTweet untuk ASC [11]. Keunggulan

ABSA dibandingkan analisis sentimen global adalah kemampuannya menghasilkan rekomendasi yang lebih terarah dan *actionable* bagi pemangku kepentingan.

Knowledge Discovery in Database (KDD)

KDD merupakan proses sistematis untuk menemukan pola dari data besar melalui lima tahap: seleksi data, pembersihan, transformasi, *data mining*, dan evaluasi hasil [12]. Kerangka ini dipilih karena memastikan setiap langkah pengolahan data memiliki keterkaitan logis dengan hasil akhir, meningkatkan reproduktibilitas dan validitas penelitian. Dalam penelitian ini KDD diterapkan mulai dari pengumpulan cuitan, *preprocessing* teks, ekstraksi aspek, klasifikasi sentimen, hingga interpretasi hasil dan penyusunan rekomendasi strategis.

IndoBERTweet

Bidirectional Encoder Representations from Transformers (BERT) memahami konteks kata secara *bidirectional*, menghasilkan representasi semantik yang lebih kaya dibandingkan model sekuensial [15]. IndoBERTweet merupakan varian BERT yang dikembangkan khusus menggunakan jutaan cuitan publik berbahasa Indonesia [6]. Model ini unggul dalam mengenali gaya bahasa informal, singkatan (*gak, bgt, utk*), *code-mixing* Bahasa Indonesia–Inggris, dan ekspresi slang lokal khas platform X. Penelitian Koto et al. (2021) menunjukkan IndoBERTweet secara konsisten mengungguli model multibahasa maupun IndoBERT generik pada tugas klasifikasi teks Twitter berbahasa Indonesia [6].

Penelitian Terdahulu

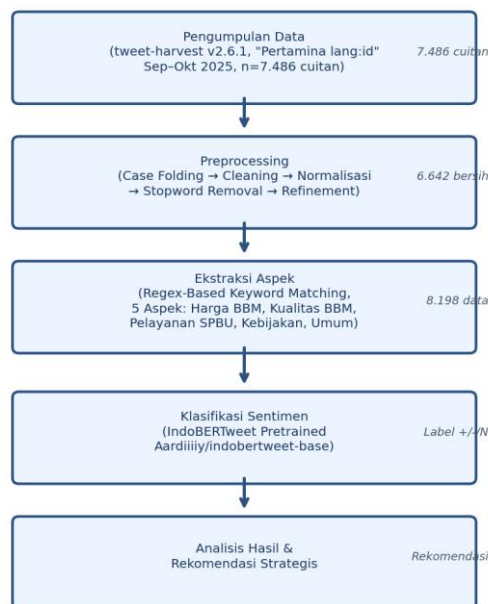
Tabel 1. Penelitian Terdahulu

| No | Penulis | Judul | Metode | Hasil Utama |
|----|-------------------------|-------------------------------|------------------------|---|
| 1 | Azhar & Khodra (2021) | ABSA Ulasan Hotel Indonesia | m-BERT Fine-tuning | Transformer lebih akurat dari LSTM pada aspek fasilitas dan pelayanan |
| 2 | Koto et al. (2021) | IndoBERTweet | IndoBERTweet | Unggul memahami bahasa informal dan singkatan Twitter Indonesia |
| 3 | Af'idah (2023) | ABSA Destinasi Wisata | LSTM/Bi-LSTM+ABSA | Transformer masih lebih unggul dalam konteks semantik |
| 4 | Salsabila et al. (2023) | Sentimen Tweet COVID-19 | ABSA+Emotion Detection | Kombinasi sentimen dan emosi memberikan pemahaman lebih lengkap |
| 5 | Sejati et al. (2024) | Sentimen Kebijakan Pemerintah | ABSA+IndoBERT | Efektif memahami opini publik secara lebih detail per aspek |

3. Metode Penelitian

Objek dan Rancangan Penelitian

Objek penelitian adalah cuitan publik di platform X berbahasa Indonesia yang membahas Pertamina pada periode September–Oktober 2025. Penelitian menggunakan pendekatan deskriptif dengan kerangka KDD, diterapkan melalui lima tahapan sistematis sebagaimana ditunjukkan pada Gambar 1.



Gambar 1. Rancangan Penelitian Menggunakan Kerangka KDD

Pengumpulan Data

Data dikumpulkan menggunakan *tweet-harvest* v2.6.1 (berbasis Node.js) dengan kata kunci "*Pertamina lang:id*" dan filter periode *since:2025-09-01 until:2025-10-31*. Autentikasi menggunakan parameter *TWITTER_AUTH_TOKEN*. Tab *LATEST* digunakan dengan batas 10.000 cuitan. Hasil disimpan dalam format CSV dengan atribut: *full_text*, *created_at*, *username*, *lang*, *favorite_count*, dan *retweet_count*. Proses ini menghasilkan 7.486 cuitan mentah.

Preprocessing

Preprocessing dilaksanakan dalam lima tahap: (1) **Case Folding** – konversi teks ke huruf kecil; (2) **Cleaning** – menghapus URL, mention, hashtag, angka, emoji, dan karakter khusus; (3) **Normalisasi** – mengubah kata tidak baku ke bentuk standar ("*gk*" → "*gak*", "*thn*" → "*tahun*"); (4) **Stopword Removal** – menghapus kata fungsi dan kata informal khas Twitter (*aja*, *sih*, *dong*, *nih*, *yg*) menggunakan daftar kustom 312 stopwords; (5) **Refinement** – menghapus duplikat dan cuitan kosong setelah pembersihan.

Ekstraksi Aspek

Ekstraksi aspek menggunakan *Regex-Based Keyword Matching* dengan lima aspek: (1) **Harga BBM** – kata kunci: harga, naik, turun, pertalite, pertamax, solar; (2) **Kualitas BBM** – kualitas, oktan, campuran, rusak, bagus, jelek, brebet; (3) **Pelayanan SPBU** – spbu, antre, antri, ramah, petugas, pelayanan; (4) **Kebijakan Pertamina** – kebijakan, pemerintah, distribusi, regulasi, subsidi, privatisasi; (5) **Umum** – cuitan yang tidak terklasifikasi ke keempat aspek. Cuitan multi-aspek direplikasi sehingga setiap baris merepresentasikan satu pasangan (aspek, sentimen), menghasilkan 8.198 baris dari 6.642 cuitan bersih [11].

Klasifikasi Sentimen

Klasifikasi sentimen menggunakan model *Aardiiiy/indobertweet-base-Indonesian-sentiment-analysis* [16] melalui pipeline *Sentiment Analysis* dari pustaka Hugging Face Transformers. Setiap cuitan diproses melalui tokenisasi, representasi vektor kontekstual, dan lapisan klasifikasi yang menghasilkan probabilitas tiga kelas. Label sentimen ditentukan melalui *argmax* probabilitas. Karena penelitian bersifat deskriptif-analitis tanpa data berlabel manual sebagai *ground truth*, evaluasi metrik performa tidak dilakukan; fokus penelitian adalah distribusi sentimen per aspek sebagai basis analisis opini publik.

4. Hasil dan Pembahasan

Hasil Preprocessing

Preprocessing mereduksi dataset dari 7.486 cuitan mentah menjadi 6.642 cuitan bersih, dengan 844 cuitan (11,27%) tereliminasi. Tahap normalisasi mengidentifikasi 2.347 kata tidak baku unik yang berhasil dinormalisasi, sementara *stopword removal* menggunakan 312 stopwords kustom. Ringkasan hasil preprocessing disajikan pada Tabel 2.

Tabel 2. Ringkasan Hasil Preprocessing Data

| Tahap Preprocessing | Jumlah Data |
|--|--------------|
| Data Mentah (Hasil Crawling) | 7.486 |
| Setelah Case Folding dan Cleaning | 7.012 |
| Setelah Normalisasi dan Stopword Removal | 6.753 |
| Data Bersih Final (Setelah Refinement) | 6.642 |
| Data Tereliminasi | 844 (11,27%) |

Hasil Ekstraksi Aspek

Proses ekstraksi aspek menghasilkan 8.198 baris data dari 6.642 cuitan, dengan 1.556 cuitan terdeteksi mengandung lebih dari satu aspek. Aspek Pelayanan SPBU mendominasi dengan 47,0% (3.856 data), mencerminkan tingginya perhatian publik terhadap pengalaman langsung di SPBU [11]. Aspek Kebijakan Pertamina didominasi kata *swasta* dan *bahlil*, mencerminkan perdebatan privatisasi dan peran Menteri ESDM. Aspek Kualitas BBM didominasi kata *rusak*, *jelek*, dan *brevet*, mengindikasikan keluhan performa kendaraan. Distribusi per aspek disajikan pada Tabel 3.

Tabel 3. Distribusi Data Berdasarkan Aspek

| Aspek | Jumlah Data | Persentase (%) |
|---------------------|-------------|----------------|
| Pelayanan SPBU | 3.856 | 47,0% |
| Umum | 2.091 | 25,5% |
| Kebijakan Pertamina | 876 | 10,7% |
| Harga BBM | 740 | 9,0% |
| Kualitas BBM | 635 | 7,7% |
| Total | 8.198 | 100,0% |

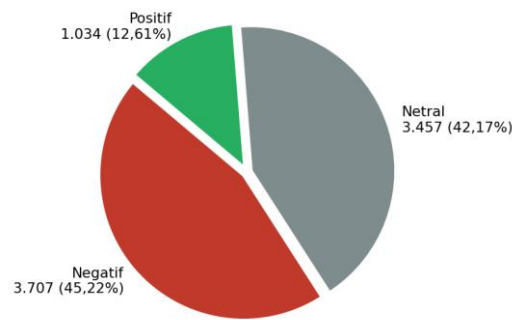
Hasil Klasifikasi Sentimen

Distribusi Sentimen Keseluruhan

Hasil klasifikasi menunjukkan sentimen Negatif mendominasi 45,22% (3.707 data), Netral 42,17% (3.457 data), dan Positif 12,61% (1.034 data). Dominasi sentimen negatif mengindikasikan opini publik yang cenderung kritis terhadap Pertamina selama September–Oktober 2025 (Tabel 4 dan Gambar 3). Hal ini sejalan dengan *negativity bias* karakteristik media sosial, di mana pengguna lebih terdorong membagikan ketidakpuasan terhadap layanan publik yang berdampak langsung pada kehidupan sehari-hari [8].

Tabel 4. Distribusi Sentimen Keseluruhan

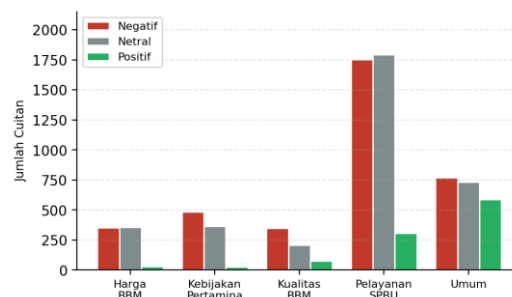
| Sentimen | Jumlah Data | Persentase (%) |
|----------|-------------|----------------|
| Negatif | 3.707 | 45,22% |
| Netral | 3.457 | 42,17% |
| Positif | 1.034 | 12,61% |
| Total | 8.198 | 100,00% |



Gambar 3. Proporsi Sentimen Keseluruhan terhadap Pertamina

Distribusi Sentimen per Aspek

Analisis sentimen per aspek mengungkap pola yang lebih informatif (Tabel 5 dan Gambar 2). Aspek **Kebijakan Pertamina** mencatat negatif tertinggi (55,4%) – isu privatisasi, distribusi subsidi, dan komunikasi kebijakan menjadi sumber kritik dominan. Aspek **Kualitas BBM** juga tinggi (55,0%) – keluhan performa kendaraan bersifat emosional karena berkaitan langsung dengan keselamatan dan biaya perawatan. Aspek **Harga BBM** menunjukkan ambivalensi publik (negatif 47,7% vs netral 48,1%) – masyarakat menyadari faktor eksternal namun tetap merasakan dampak daya beli. Aspek **Pelayanan SPBU** menjadi satu-satunya di mana netral (47,3%) sedikit melampaui negatif (46,6%), mengindikasikan pengalaman beragam bergantung lokasi SPBU. Aspek **Umum** memiliki distribusi paling berimbang dengan positif tertinggi (28,2%), mencerminkan persepsi keseluruhan yang lebih baik terhadap Pertamina sebagai perusahaan energi nasional.



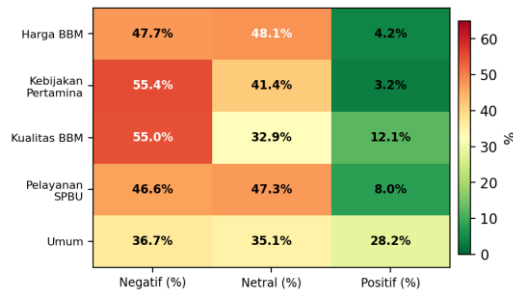
Gambar 2. Distribusi Sentimen per Aspek

Tabel 5. Distribusi Sentimen per Aspek

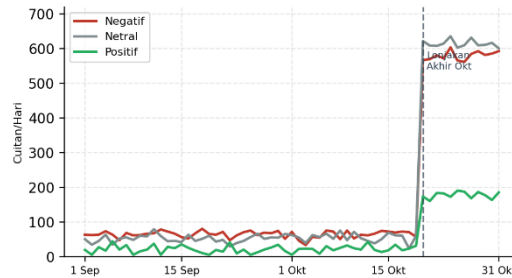
| Aspek | Negatif | Netral | Positif |
|---------------------|----------------|----------------|----------------|
| Harga BBM | 353 (47,7%) | 356 (48,1%) | 31 (4,2%) |
| Kebijakan Pertamina | 485 (55,4%) | 363 (41,4%) | 28 (3,2%) |
| Kualitas BBM | 349 (55,0%) | 209 (32,9%) | 77 (12,1%) |
| Pelayanan SPBU | 1.752 (46,6%) | 1.795 (47,3%) | 309 (8,0%) |
| Umum | 768 (36,7%) | 734 (35,1%) | 589 (28,2%) |
| Total | 3.707 (45,22%) | 3.457 (42,17%) | 1.034 (12,61%) |

Analisis Tren Temporal

Heatmap (Gambar 4) memberikan visualisasi komparatif efisien; gradasi merah pada Kebijakan Pertamina dan Kualitas BBM mengonfirmasi kedua aspek sebagai paling kritis. Analisis tren temporal (Gambar 5) mengungkap pola signifikan: seluruh aspek menunjukkan volume stabil selama September hingga pertengahan Oktober 2025, kemudian lonjakan tajam pada akhir Oktober. Sentimen negatif mencapai puncak ~580 cuitan/hari dan netral ~620 cuitan/hari. Lonjakan simultan lintas aspek mengindikasikan adanya peristiwa spesifik yang memicu reaksi publik masif – penting sebagai sinyal deteksi dini bagi sistem pemantauan sentimen Pertamina.



Gambar 4. Heatmap Persentase Sentimen per Aspek (%)



Gambar 5. Tren Sentimen dari Waktu ke Waktu (Sep–Okt 2025)

Rekomendasi Strategis

Berdasarkan distribusi sentimen per aspek, tiga prioritas strategis direkomendasikan:

(1) **Transparansi Komunikasi Kebijakan:** Mengingat proporsi negatif tertinggi pada Kebijakan (55,4%), Pertamina perlu menjadikan platform X sebagai saluran komunikasi *two-way* aktif: menyampaikan klarifikasi resmi, menjelaskan mekanisme penetapan harga melalui infografis berkala, dan merespons keluhan publik secara cepat berbasis data.

(2) **Perbaikan Kualitas dan Verifikasi Produk:** Untuk Kualitas BBM (negatif 55,0%), peningkatan frekuensi pengujian kualitas independen dengan publikasi terbuka, serta program edukasi konsumen mengenai standar oktan, dapat membangun kepercayaan dan mengurangi miskomunikasi.

(3) **Standarisasi Layanan SPBU:** Aspek Pelayanan SPBU memiliki volume tertinggi (47,0%). Program pelatihan petugas secara nasional, optimalisasi sistem antrian digital, kalibrasi alat ukur, dan kanal pengaduan *real-time* merupakan langkah konkret yang direkomendasikan.

5. Kesimpulan

Penelitian ini berhasil menerapkan ABSA menggunakan model IndoBERTtweet *pretrained* untuk menganalisis opini publik terhadap Pertamina di platform X periode September–Oktober 2025. Dari 6.642 cuitan bersih diperoleh 8.198 data setelah ekstraksi lima aspek: Pelayanan SPBU (47,0%), Umum (25,5%), Kebijakan Pertamina (10,7%), Harga BBM (9,0%), dan Kualitas BBM (7,7%). Sentimen Negatif mendominasi (45,22%), diikuti Netral (42,17%) dan Positif (12,61%). Kebijakan Pertamina (55,4%) dan Kualitas BBM (55,0%) mencatat proporsi negatif tertinggi; aspek Umum menunjukkan distribusi paling berimbang dengan positif tertinggi (28,2%).

Tiga rekomendasi strategis utama: (1) peningkatan transparansi komunikasi kebijakan harga dan regulasi energi; (2) penguatan program verifikasi dan publikasi kualitas BBM secara independen; (3) standarisasi layanan SPBU nasional melalui pelatihan, teknologi antrian digital, dan kanal pengaduan *real-time*. Keterbatasan penelitian mencakup tidak adanya *ground truth* berlabel manual dan cakupan satu platform. Penelitian selanjutnya disarankan menambahkan anotasi manual, memperluas ke platform lain, dan mengembangkan ekstraksi aspek berbasis NER.

Daftar Pustaka

- [1] F. Koto, J. H. Lau, and T. Baldwin, "IndoBERTtweet: A pretrained language model for Indonesian Twitter," in Proc. EMNLP 2021, 2021.
- [2] DataReportal, "Digital 2023: Indonesia," 2023. [Online]. Available: <https://datareportal.com/reports/digital-2023-indonesia>
- [3] R. Halim, Y. Nugroho, and P. I. Santosa, "Sentiment analysis of Indonesian tweets using Bidirectional LSTM," *Procedia Comput. Sci.*, vol. 179, pp. 476–484, 2021.
- [4] M. N. Azhar and M. L. Khodra, "Aspect-based sentiment analysis for Indonesian hotel reviews using BERT," *Procedia Comput. Sci.*, vol. 179, pp. 455–463, 2021.
- [5] N. Said and T. Manik, "Analyzing political sentiment in Indonesian elections using ABSA," *Procedia Comput. Sci.*, vol. 217, pp. 124–132, 2023.
- [6] F. Koto, J. H. Lau, and T. Baldwin, "IndoBERTtweet," Proc. EMNLP 2021, 2021.

- [7] A. Kurniawan, M. F. Pratama, and S. Wibowo, “Aspect-based sentiment analysis for Indonesian e-commerce using BERT and CNN,” *J. Inf. Syst. Informatics*, vol. 4, no. 3, pp. 567–579, 2022.
- [8] B. Liu, *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions*, 2nd ed. Cambridge Univ. Press, 2020.
- [9] M. Pontiki et al., “SemEval-2016 Task 5: Aspect based sentiment analysis,” *Proc. SemEval-2016*, pp. 19–30, 2016.
- [10] V. Giudice and V. Cutello, “European social sentiment during energy crises: A neural transformer approach,” *Inf. Sci.*, vol. 578, pp. 786–799, 2021.
- [11] H. Sutanto and I. Darmawan, “Aspect extraction and sentiment mapping in Indonesian transportation reviews,” *JITEKI*, vol. 9, no. 2, pp. 224–234, 2023.
- [12] R. Sudarsono and N. Fitriani, “Implementasi KDD untuk Analisis Data Media Sosial,” *J. Teknol. Inf. Ilmu Komput.*, vol. 8, no. 3, pp. 411–420, 2021.
- [13] M. Handayani and F. Sudibyoy, “Tracking social media sentiment toward fuel policy in Indonesia,” *Sustainability*, vol. 14, no. 17, p. 10891, 2022.
- [14] M. Hidayat and S. Yuliana, “Preprocessing and normalization methods for Indonesian social media sentiment analysis,” *IJCCS*, vol. 16, no. 2, pp. 112–121, 2022.
- [15] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of deep bidirectional transformers,” *Proc. NAACL-HLT 2019*, pp. 4171–4186, 2019.
- [16] M. Ardiyanto, “Aardiiiy/indobertweet-base-Indonesian-sentiment-analysis,” Hugging Face, 2024. [Online]. Available: <https://huggingface.co/Aardiiiy>