

ANALISIS SENTIMEN PUBLIK TERHADAP KASUS KERACUNAN MAKANAN MBG BERBASIS DATA MEDIA SOSIAL MENGGUNAKAN MODEL INDOBERT

Alustina Suci Manah¹, Riza Ibnu Adam², Aries Suharso³

^{1,2,3} Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Singaperbangsa Karawang; Jl. H.S. Ronggowaluyo, Puseurjaya, Telukjambe Timur, Karawang, Jawa Barat 41361; Telp. (0267) 641177.

Keywords:

Analisis Sentimen;
Keracunan Makanan;
IndoBERT; MBG; Platform X.

Correspondent Email:

2210631170006@student.unsika.ac.id, @unsika.ac.id

Abstrak. Penelitian ini dilatarbelakangi oleh meluasnya kekhawatiran publik di media sosial terhadap insiden keracunan makanan dalam pelaksanaan Program Makan Bergizi Gratis (MBG), yang memicu krisis kepercayaan terhadap keamanan program. Penelitian ini bertujuan memetakan sentimen masyarakat di Platform X dan menguji efektivitas model IndoBERT-Base-P2 dalam mengklasifikasikan opini publik berbahasa Indonesia. Metode yang digunakan adalah *Knowledge Discovery in Database* (KDD) dengan 5.082 cuitan periode Januari–Desember 2025. Tahap *pre-processing* menggunakan kamus normalisasi kustom 2.127 entri tanpa *stemming* dan *stopword removal*. Pelabelan sentimen dilakukan melalui pendekatan hibrida (IndoBERT & *Profanity Rule*), sementara *fine-tuning* menggunakan IndoBERT-Base-P2 sebagai base model *pre-trained* murni. *Hyperparameter* optimal diperoleh dari *full grid search* 12 kombinasi dengan konfigurasi terbaik: *Max Length* 128, *Batch Size* 16, *Learning Rate* 3×10^{-5} , *best epoch* ke-7 pada rasio 60:20:20. Hasil pengujian menunjukkan akurasi 94,75% dan *F1-Score* tertimbang 0,95. Distribusi sentimen didominasi sentimen negatif 74,27% (3.325 cuitan), netral 16,51% (739 cuitan), dan positif 9,22% (413 cuitan). Temuan ini membuktikan insiden keracunan MBG memicu penolakan masif di media sosial dan model IndoBERT-Base-P2 terbukti efektif sebagai alat pemetaan opini publik untuk evaluasi kebijakan pemerintah.



Copyright © [JITET](http://www.jitet.org) (Jurnal Informatika dan Teknik Elektro Terapan). This article is an open access article distributed under terms and conditions of the Creative Commons Attribution (CC BY NC)

Abstract. This research is motivated by widespread public concern on social media regarding food poisoning incidents in the implementation of the Free Nutritious Meal Program (MBG), which triggered a crisis of public confidence. This study aims to map public sentiment on Platform X and evaluate the effectiveness of the IndoBERT-Base-P2 model in classifying Indonesian public opinion. The method employed is Knowledge Discovery in Database (KDD) using 5,082 tweets from January to December 2025. Pre-processing used a custom normalization dictionary of 2,127 entries without stemming and stopword removal. Sentiment labeling was performed through a hybrid approach (IndoBERT & Profanity Rule), while fine-tuning used IndoBERT-Base-P2 as a pure pre-trained base model. The optimal configuration was obtained from a full grid search of 12 combinations, yielding Max Length 128, Batch Size 16, Learning Rate 3×10^{-5} , best epoch 7 on the 60:20:20 split. Evaluation demonstrated 94.75% accuracy and a weighted F1-Score of 0.95. Sentiment distribution was dominated by Negative at 74.27% (3,325 tweets), Neutral at 16.51%, and Positive at 9.22%. These findings confirm that MBG poisoning incidents triggered massive public rejection on social media, and the IndoBERT-Base-P2 model proved effective for public opinion mapping in support of government policy evaluation.

1. PENDAHULUAN

Perkembangan teknologi informasi dan komunikasi yang sangat pesat telah mengubah cara masyarakat berinteraksi dan memperoleh informasi. Media sosial kini menjadi salah satu ruang utama bagi masyarakat untuk menyuarkan opini, berdiskusi, bahkan membentuk persepsi terhadap berbagai isu publik dan kebijakan pemerintah. Platform seperti X (Twitter), Instagram, dan TikTok menjadi wadah aspirasi digital di mana masyarakat bebas mengekspresikan pandangan mereka terhadap isu-isu sosial, ekonomi, maupun politik [1].

Salah satu isu publik yang menjadi sorotan besar pada tahun 2025 di Indonesia adalah kasus keracunan makanan pada pelaksanaan Program Makan Bergizi Gratis (MBG). Program MBG merupakan kebijakan pemerintah yang bertujuan menyediakan makan siang gratis bagi peserta didik untuk meningkatkan gizi dan kualitas sumber daya manusia. Meskipun kebijakan ini memiliki tujuan positif, pelaksanaannya sempat memunculkan berbagai persoalan, terutama setelah munculnya beberapa laporan kasus keracunan makanan massal di sejumlah sekolah. Berdasarkan data CISDI & Riset Kumbaran (2025), tercatat total 6.123 kasus dengan puncak tertinggi pada bulan September 2025 sebanyak 2.223 kasus [2].

Melalui analisis data media sosial, peneliti dan pembuat kebijakan dapat memperoleh pemahaman yang lebih mendalam mengenai persepsi masyarakat terhadap suatu isu secara real-time. Namun, karena data yang dihasilkan berbentuk teks tidak terstruktur dan jumlahnya sangat besar, maka diperlukan metode komputasional yang mampu mengekstrak informasi secara sistematis dan akurat, salah satunya melalui analisis sentimen [3].

Pemilihan model klasifikasi teks menjadi aspek yang sangat penting dalam memperoleh hasil analisis yang akurat. Salah satu model yang terbukti unggul dalam memahami konteks Bahasa Indonesia adalah IndoBERT, yaitu model transformer-based yang dikembangkan dari arsitektur BERT (Bidirectional Encoder Representations from Transformers) dan dilatih khusus menggunakan korpus Bahasa Indonesia [4]. Penelitian oleh Lukmanto & Wijayanti

(2025) menunjukkan bahwa model IndoBERT-Large-P2 menghasilkan akurasi sebesar 80,19% dan AUC rata-rata 91% dalam menganalisis sentimen program MBG [5]. Penelitian Manoppo et al. (2025) juga menunjukkan performa unggul model IndoBERT dengan akurasi 84,94% pada konteks kebijakan kenaikan PPN 12% [6].

Berbeda dengan penelitian terdahulu, penelitian ini memiliki fokus yang lebih spesifik pada kasus keracunan makanan dalam pelaksanaan MBG, sebuah isu aktual yang memicu reaksi publik luas dan fluktuasi sentimen yang signifikan. Penelitian ini menerapkan pendekatan hybrid labeling yang memisahkan peran dua model berbeda untuk menghindari circular bias, mengeksplorasi variasi jumlah epoch (1-10) dan tiga variasi rasio pembagian data (80:10:10, 70:15:15, 60:20:20), serta menggunakan kamus normalisasi kustom 2.127 entri. Berdasarkan uraian tersebut, penelitian ini bertujuan untuk: (1) menganalisis persepsi dan sentimen publik terhadap kasus keracunan makanan MBG di media sosial X; (2) menerapkan model IndoBERT-Base-P2 dalam klasifikasi sentimen; dan (3) mengevaluasi pengaruh variasi jumlah epoch terhadap performa model.

2. TINJAUAN PUSTAKA

2.1 Analisis Sentimen

Analisis sentimen merupakan proses sistematis untuk mengidentifikasi, mengekstraksi, dan mengevaluasi sikap, emosi, atau opini yang terkandung dalam sebuah teks. Teknik ini merupakan cabang penting dalam bidang *Natural Language Processing* (NLP) yang bertujuan untuk memahami persepsi publik terhadap suatu entitas, seperti produk, kebijakan, layanan, maupun peristiwa sosial [7]. Secara umum, analisis sentimen dapat dilakukan melalui tiga pendekatan: (1) berbasis kamus (*lexicon-based*), (2) berbasis pembelajaran mesin (*machine learning*), dan (3) pendekatan hibrid yang menggabungkan keduanya.

2.2 Knowledge Discovery in Databases (KDD)

KDD merupakan proses sistematis untuk mengekstraksi pengetahuan atau pola yang

bermanfaat dari kumpulan data besar. Menurut Collins et al. (2021), KDD terdiri atas serangkaian tahapan yang saling berhubungan untuk mengubah data mentah menjadi informasi yang bernilai [8]. Tahapan KDD meliputi: *Data Selection*, *Pre-processing*, *Transformation*, *Data Mining*, dan *Evaluation and Interpretation*.

2.3 Model BERT dan IndoBERT

BERT (*Bidirectional Encoder Representations from Transformers*) adalah model representasi bahasa yang diperkenalkan oleh Google AI pada tahun 2019, menggunakan arsitektur Transformer dengan mekanisme *self-attention* untuk memproses konteks kata secara dua arah [9]. IndoBERT adalah adaptasi monolingual dari BERT khusus untuk Bahasa Indonesia yang dikembangkan untuk mengisi kekurangan model bahasa berbahasa Indonesia. Model ini dilatih menggunakan korpus besar Bahasa Indonesia, mencakup Wikipedia Indonesia dan korpus web lokal [4]. *Fine-tuning* merupakan tahap pelatihan lanjutan yang dilakukan setelah model bahasa pra-latih disesuaikan terhadap tugas tertentu, memungkinkan model diadaptasi ke tugas klasifikasi sentimen spesifik dengan menambahkan lapisan output di bagian akhir arsitektur.

3. METODE PENELITIAN

Penelitian ini menggunakan metode *Knowledge Discovery in Database* (KDD) yang mencakup lima tahapan sistematis: *Data Selection*, *Preprocessing*, *Transformation*, *Data Mining*, dan *Evaluation*. Objek penelitian adalah cuitan dari Platform X (Twitter) yang membahas topik kasus keracunan makanan dalam Program Makan Bergizi Gratis (MBG) pada rentang waktu Januari hingga Desember 2025, dengan bahasa Indonesia.

3.1 Data Selection

Pengumpulan data dilakukan melalui teknik *scraping* dari Platform X menggunakan tools *Tweet-Harvest* dalam lingkungan Google Colaboratory dengan bahasa pemrograman Python. Data dikumpulkan menggunakan tiga strategi: tagar (#keracunanmbg), postingan akun media tertentu, dan kata kunci "Keracunan MBG", menghasilkan total 5.742 data mentah. Setelah penghapusan 495 data duplikat dan 165 data *garbage* melalui validasi manual,

diperoleh 5.082 data valid yang siap diproses lebih lanjut.

3.2 Preprocessing

Tahap *preprocessing* terdiri dari empat langkah berurutan: (1) *Cleansing* — menghapus URL, *mention*, tagar, emoji, dan karakter non-alfabet; (2) *Case Folding* — mengonversi teks menjadi huruf kecil; (3) Normalisasi — mengubah kata tidak baku dan *slang* menggunakan kamus normalisasi kustom 2.127 entri (kamus_normalisasi_mbg.json); dan (4) Tokenisasi menggunakan fungsi *word_tokenize* dari NLTK. Penelitian ini secara sengaja tidak melakukan *stemming* dan *stopword removal* karena arsitektur *bidirectional* IndoBERT memerlukan konteks kalimat utuh untuk memahami emosi pengguna secara presisi [10].

3.3 Transformation

Tahap *Transformation* mencakup dua proses utama. Pertama, pelabelan data menggunakan *Hybrid Labeling* yang memisahkan secara tegas peran dua model: model *mdhugol/indonesia-bert-sentiment-classification* digunakan sebagai alat pelabelan otomatis, sedangkan model *indobenchmark/indobert-base-p2* digunakan sebagai *base model fine-tuning*. Pemisahan ini untuk menghindari *circular bias*. *Profanity rule* berbasis leksikon (18 kata kasar) diterapkan sebagai lapisan koreksi, berhasil mengoreksi 9 data. Setelah filtrasi *confidence threshold* $\geq 0,75$, diperoleh 4.477 data berlabel. Kedua, pembagian data dilakukan dengan tiga variasi rasio: 80:10:10, 70:15:15, dan 60:20:20, menggunakan parameter *stratify* untuk menjaga proporsionalitas distribusi kelas.

3.4 Data Mining

Tahap *Data Mining* menggunakan model *IndoBERT-Base-P2* (*indobenchmark/indobert-base-p2*) sebagai *base model fine-tuning*. Penentuan konfigurasi *hyperparameter* optimal dilakukan melalui eksperimen *full grid search* terhadap 12 kombinasi dengan variasi *Max Length* (64 dan 128), *Batch Size* (16 dan 32), dan *Learning Rate* (2×10^{-5} dan 3×10^{-5}). Selanjutnya, variasi jumlah *epoch* (1 hingga 10) diuji pada setiap variasi rasio *splitting* untuk menentukan konfigurasi pelatihan terbaik. Mekanisme *best model saving* berdasarkan

validation accuracy tertinggi diterapkan untuk mencegah *overfitting*.

3.5 Evaluation

Evaluasi performa model dilakukan menggunakan metrik *Confusion Matrix* yang mencakup *accuracy*, *precision*, *recall*, dan *F1-score* untuk setiap kelas sentimen (positif, negatif, netral). Evaluasi akhir dilakukan pada data uji yang belum pernah dilihat model selama pelatihan.

4. HASIL DAN PEMBAHASAN

4.1 Data Selection

Pengumpulan data menghasilkan 5.742 data mentah dari tiga strategi: kata kunci "Keracunan MBG" menghasilkan 4.643 data (80,9%), tagar #keracunanmbg menghasilkan 768 data, dan strategi berbasis akun menghasilkan 331 data dari tiga akun media. Setelah dua tahap filtrasi — penghapusan 495 data duplikat (8,6%) dan 165 data *garbage* (3,1%) melalui validasi manual — diperoleh 5.082 data valid (96,9%) sebagai dataset final.

Tabel 1. Rekapitulasi Tahapan Data Selection

Tahapan	Proses	Jumlah Data
Scraping & Penggabungan	3 strategi paralel	5.742
Penghapusan Duplikat	<i>drop_duplicates()</i>	5.247 (-495)
Validasi Manual (Garbage)	<i>Filter is_garbage = 0</i>	5.082 (-165)
Dataset Final	<i>Siap Preprocessing</i>	5.082

4.2 Preprocessing

Tahap *preprocessing* berhasil membersihkan 5.082 data dari berbagai *noise* khas platform X seperti URL, *mention*, tagar, dan emoji melalui proses *cleansing* menggunakan modul *Regular Expression* (re) Python. Penyeragaman huruf melalui *case folding* memastikan konsistensi representasi kata. Normalisasi menggunakan kamus kustom 2.127 entri berhasil mengubah singkatan informal seperti "mbg" menjadi "makan bergizi gratis", "gk" menjadi "tidak", dan "yg" menjadi "yang", sehingga model dapat mengenali entitas semantik secara konsisten. Tokenisasi menggunakan NLTK menghasilkan 5.082 data

bertokenisasi yang siap untuk tahap *Transformation*.

4.3 Transformation dan Distribusi Sentimen

Proses *Hybrid Labeling* berhasil melabeli 5.072 data menggunakan model *mdhugol* sebagai *labeler* otomatis yang dikombinasikan dengan *profanity rule* (18 kata kasar). Rule berbasis leksikon mengoreksi 9 data yang awalnya diprediksi bukan negatif. Setelah filtrasi *confidence threshold* $\geq 0,75$, diperoleh 4.477 data berlabel final. Distribusi label menunjukkan ketimpangan kelas yang mencerminkan realita opini publik terhadap kasus keracunan MBG.

Tabel 2. Distribusi Label Sentimen Final

Kelas Sentimen	Jumlah Data	Persentase
Negatif	3.325	74,27%
Netral	739	16,51%
Positif	413	9,22%
Total	4.477	100%

4.4 Data Mining: Konfigurasi Hyperparameter

Eksperimen *full grid search* terhadap 12 kombinasi *hyperparameter* menghasilkan konfigurasi optimal: *Max Length* 128, *Batch Size* 16, dan *Learning Rate* 3×10^{-5} dengan *validation accuracy* 95,31%. Konfigurasi ini berbeda dari beberapa penelitian serupa yang menggunakan *Max Length* 64, hal ini dapat dijelaskan oleh karakteristik IndoBERT-P2 sebagai model *pre-trained* murni yang memerlukan *max length* lebih panjang untuk menangkap konteks yang lebih kaya.

Tabel 3. Perbandingan Best Epoch Antar Variasi Split

Rasio Split	Best Epoch	Val Accuracy	Test Accuracy
80:10:10	6	96,21%	93,97%
70:15:15	8	95,09%	94,49%
60:20:20	7	95,31%	94,75%

Eksperimen variasi *epoch* (1 hingga 10) dan tiga variasi rasio *splitting* memberikan gambaran komprehensif tentang dinamika pelatihan IndoBERT-P2. Sebagai model *pre-trained* murni, IndoBERT-P2 memerlukan beberapa *epoch* awal untuk membangun

pemahaman dasar tentang pola sentimen sebelum konvergensi tercapai, dengan lonjakan akurasi terbesar terjadi pada *Epoch 2* di semua variasi split. Setelah *best epoch*, model menunjukkan gejala *overfitting* yang ditandai oleh *training loss* mendekati nol sementara *validation accuracy* stagnan, memvalidasi pentingnya mekanisme *best model saving*.

4.5 Evaluation

Pengujian final pada 896 data uji (split 60:20:20) menghasilkan akurasi akhir 94,75% (848 prediksi benar dari 896 total data uji). Laporan klasifikasi per kelas disajikan pada Tabel 4.

Tabel 4. Laporan Klasifikasi Final Model IndoBERT-Base-P2

Kelas	Precision	Recall	F1-Score	Support
Positif	0,81	0,76	0,78	83
Netral	0,93	0,95	0,94	148
Negatif	0,97	0,97	0,97	665
<i>Weighted Average</i>	0,95	0,95	0,95	896

Kelas Negatif memiliki performa tertinggi dengan *F1-Score* 0,97, mengkonfirmasi kemampuan model dalam mendeteksi keluhan dan kritik masyarakat secara akurat — ini merupakan pencapaian paling penting mengingat kelas negatif adalah fokus utama analisis sentimen terhadap kasus MBG. Kelas Netral menunjukkan performa sangat baik dengan *F1-Score* 0,94, menandakan model mampu membedakan cuitan informatif faktual dari cuitan emosional dengan presisi tinggi. Kelas Positif mencatat *F1-Score* 0,78, lebih rendah dibandingkan dua kelas lainnya akibat keterbatasan jumlah data latih pada kelas minoritas (hanya 9,22%).

Perbandingan antar variasi split menunjukkan tren yang konsisten: split 60:20:20 menghasilkan *test accuracy* tertinggi (94,75%) dibandingkan 70:15:15 (94,49%) dan 80:10:10 (93,97%). Temuan ini mengindikasikan bahwa estimasi performa pada data uji yang lebih besar lebih andal dan representatif terhadap kemampuan generalisasi model sesungguhnya. *Test accuracy* sebesar 94,75% melampaui penelitian Lukmanto & Wijayanti (2025) yang menggunakan IndoBERT-Large-P2 (akurasi 80,19%) dan Manoppo et al. (2025) yang mencapai 84,94%,

mengkonfirmasi efektivitas pendekatan metodologis yang digunakan.

4.6 Analisis Distribusi Sentimen

Klasifikasi pada seluruh dataset mengungkap ketimpangan yang sangat signifikan: Sentimen Negatif mendominasi secara mutlak sebesar 74,27% (3.325 data), disusul sentimen Netral sebesar 16,51% (739 data), dan sentimen Positif sebesar 9,22% (413 data). Dominasi penolakan yang mencapai hampir tiga perempat total data membuktikan bahwa insiden keracunan makanan MBG telah memicu kekhawatiran dan krisis kepercayaan publik yang masif.

Analisis *Word Cloud* pada masing-masing kelompok sentimen mengungkapkan narasi yang berbeda. Kelompok sentimen Negatif menampilkan kata dominan seperti "bencana", "nyawa", "rakyat", dan "anggaran", menggambarkan bahwa publik tidak memandang insiden keracunan sebagai kecelakaan teknis biasa, melainkan sebagai kegagalan sistemik yang mengancam kredibilitas program. Kelompok sentimen Netral didominasi kata "satuan", "pelayanan", "pemenuhan", dan "diduga", bertindak sebagai penyebar informasi faktual. Kelompok sentimen Positif yang paling sedikit menampilkan kata "nyata", "aman", "semoga", dan "sukses", mencerminkan aspirasi konstruktif untuk perbaikan program di masa depan.

5. KESIMPULAN

- a. Persepsi dan sentimen publik terhadap kasus keracunan makanan MBG di Platform X didominasi oleh sentimen negatif yang sangat masif sebesar 74,27% (3.325 cuitan), mencerminkan krisis kepercayaan publik yang kuat terhadap pelaksanaan program. Sentimen netral (16,51%) berperan sebagai penyebar informasi faktual, sedangkan sentimen positif (9,22%) mencerminkan harapan akan perbaikan kualitas program.
- b. Penerapan model IndoBERT-Base-P2 terbukti efektif dalam mengklasifikasikan sentimen publik yang kompleks, mencapai *test accuracy* 94,75% dan *F1-Score* rata-rata tertimbang 0,95 pada konfigurasi

terbaik (split 60:20:20, best epoch ke-7). Pendekatan pemisahan peran antara model labeling dan model fine-tuning berhasil menghindari circular bias. Kelas Negatif mencatatkan performa tertinggi dengan F1-Score 0,97.

- c. Variasi jumlah epoch (1 hingga 10) memiliki pengaruh signifikan terhadap performa model. Best epoch bervariasi antara 6 hingga 8 tergantung rasio splitting, dengan rasio 60:20:20 menghasilkan estimasi performa paling andal. Secara keseluruhan, temuan penelitian ini membuktikan bahwa keresahan masyarakat terhadap kasus keracunan MBG dapat diukur secara ilmiah, dan hasil analisis ini dapat dijadikan rujukan berbasis data bagi pemerintah untuk melakukan evaluasi terhadap manajemen keamanan pangan Program MBG.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Bapak Riza Ibnu Adam, M.Si. selaku Pembimbing I dan Bapak Dr. Aries Suharso, S.Si., M.Kom. selaku Pembimbing II yang telah membimbing dan memberikan arahan selama penelitian ini. Ucapan terima kasih juga disampaikan kepada Fakultas Ilmu Komputer Universitas Singaperbangsa Karawang atas dukungan yang diberikan.

DAFTAR PUSTAKA

- [1] D. Latif, M. A. Samad, Rinawulandari, dan S. A. Kadir, "Social Media in Shaping Public Opinion Roles and Impact: A Systematic Review," *Jurnal Komunikasi: Malaysian Journal of Communication*, vol. 40, no. 2, pp. 205–223, 2024.
- [2] CISDI & Riset Kumparan, "Data Kasus Keracunan MBG 2025," Kumparan, 2025. [Online]. Available: <https://kumparan.com>. [Accessed: Mar. 2026].
- [3] G. P. Rizky dan W. Alrasyid, "AI-Based Sentiment Analysis of Social Media to Detect Public Opinion on Government Policies," *Journal Basic Science and Technology*, vol. 14, no. 2, pp. 61–69, 2025.
- [4] F. Koto, A. Rahimi, J. H. Lau, dan T. Baldwin, "IndoLEM and IndoBERT: A Benchmark Dataset and Pre-trained Language Model for Indonesian NLP," dalam *Proc. COLING 2020*, pp. 757–770.
- [5] J. A. Lukmanto dan T. P. Wijayanti, "Analisis Sentimen Program Makan Siang Gratis Menggunakan Model IndoBERT," *Jurnal Algoritma, Logika dan Komputasi*, vol. 8, no. 2, pp. 820–829, 2025.
- [6] M. R. Manoppo et al., "Analisis Sentimen Publik Di Media Sosial Terhadap Kenaikan PPN 12% Di Indonesia Menggunakan Indobert," *Jurnal Kecerdasan Buatan dan Teknologi Informasi*, vol. 4, no. 2, pp. 152–163, 2025.
- [7] M. A. Gumilang et al., "Sentiment Analysis of Indonesian Ministries Social Media: Citizen Responses Utilizing TextBlob Analyser," *Jurnal Socioteknologi*, vol. 23, no. 2, pp. 203–216, 2024.
- [8] C. Collins, D. Dennehy, K. Conboy, dan P. Mikalef, "Artificial intelligence in information systems research: A systematic literature review and research agenda," *Int. J. Inf. Manag.*, vol. 60, 2021.
- [9] J. Devlin, M. W. Chang, K. Lee, dan K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," dalam *Proc. NAACL HLT 2019*, pp. 4171–4186.
- [10] N. P. D. Agustina, I. D. A. P. P. Tentrinjaya, dan I. G. N. L. Wijayakusuma, "Public sentiment analysis on demonstration actions using IndoBERT based on transfer learning," *Journal of Applied Informatics and Computing (JAIC)*, vol. 10, no. 1, pp. 969–974, 2026.
- [11] M. N. Zaidan, Y. Sibaroni, dan S. S. Prasetyowati, "Learning rate and epoch optimization in the fine-tuning process for IndoBERT's performance on sentiment analysis of MyTelkomsel app reviews," *Jurnal Teknik Informatika (Jutif)*, vol. 5, no. 5, pp. 1443–1450, 2024.
- [12] F. Y. A'la, "Optimasi Klasifikasi Sentimen Ulasan Game Berbahasa Indonesia: IndoBERT dan SMOTE untuk Menangani Ketidakseimbangan Kelas," *Edumatic: Jurnal Pendidikan Informatika*, vol. 9, no. 1, pp. 256–265, 2025.
- [13] C. J. L. Tobing, I. G. N. L. Wijayakusuma, dan L. P. I. Harini, "Perbandingan Kinerja IndoBERT dan MBERT Untuk Deteksi Berita Hoaks Politik dalam Bahasa Indonesia," *JST (Jurnal Sains dan Teknologi)*, vol. 14, no. 1, pp. 114–123, 2025.
- [14] N. N. Saragih dan R. Kurniawan, "Fine-Tuning IndoBERT untuk Klasifikasi Sentimen," *Journal of Computer Engineering, System and Science*, vol. 10, no. 1, pp. 299–311, 2025.
- [15] U. Agustini dan S. Mulyani, "Efektivitas dan Tantangan Kebijakan Program Makan Bergizi Gratis sebagai Intervensi Pendidikan di

- Indonesia," Jurnal Kiprah Pendidikan, vol. 4, no. 3, pp. 362–368, 2025.
- [16] D. Duei Putri, G. F. Nama, dan W. E. Sulistiono, "Analisis Sentimen Kinerja Dewan Perwakilan Rakyat (DPR) Pada Twitter Menggunakan Metode Naive Bayes Classifier," JITET, vol. 10, no. 1, Jan. 2022.