

Analisis Sentimen Pengguna X Terhadap Kebijakan Tes Kemampuan Akademik (TKA) Jenjang SMA dan Sederajat Menggunakan Metode IndoBERT

Keisha Kayana Aptadhea^{*1}, NH Waryanto²

^{1,2}Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Yogyakarta
Email: ¹keishakayana.2022@student.uny.ac.id, ²nur_hw@uny.ac.id

Abstrak

Kebijakan Tes Kemampuan Akademik (TKA) sebagai bentuk evaluasi pendidikan memunculkan beragam respons masyarakat di media sosial yang perlu dipetakan secara sistematis. Penelitian ini bertujuan untuk mengetahui distribusi sentimen pengguna X terhadap kebijakan TKA jenjang SMA dan menganalisis kinerja model IndoBERT varian *indobert-base-p2* dalam mengklasifikasikan sentimen positif, netral, dan negatif. Data berupa 4.420 cuitan berbahasa Indonesia periode 25 Mei 2025–5 Januari 2026 dikumpulkan menggunakan Tweet-Harvest dan diproses melalui *case folding*, *cleaning*, serta normalisasi. Pelabelan dilakukan menggunakan *pseudo-labeling* dengan *confidence threshold* 0,7 dan pelabelan manual, dengan kesesuaian diukur menggunakan Cohen's Kappa sebesar 0,808. Model IndoBERT di-*fine-tune* menggunakan 70% data latih dan dievaluasi terhadap 884 data uji berlabel *gold standard*. Hasil menunjukkan sentimen negatif mendominasi sebesar 50,4%, diikuti netral 34,6% dan positif 15,0%. Model memperoleh *accuracy*, *macro F1-score*, dan *weighted F1-score* masing-masing sebesar 0,91, menunjukkan performa klasifikasi yang baik. Temuan ini dapat membantu memetakan respons publik sebagai bahan evaluasi dan komunikasi kebijakan pendidikan. Penelitian ini menyarankan perlunya sosialisasi kebijakan TKA yang lebih jelas, bertahap, dan transparan.

Kata kunci: analisis sentimen, *fine-tuning*, IndoBERT, media sosial X, tes kemampuan akademik

Sentiment Analysis Of X Users' Opinions Toward The Academic Ability Test (TKA) Policy For Senior High School Students Using IndoBERT

Abstract

The Academic Ability Test (Tes Kemampuan Akademik/TKA) as a national education evaluation policy has generated diverse public responses on social media that need to be systematically mapped. This study aims to determine the sentiment distribution of platform X users toward the TKA policy for senior high school students and to evaluate the performance of the IndoBERT model variant *indobert-base-p2* in classifying sentiments into positive, neutral, and negative categories. A total of 4,420 Indonesian-language tweets from May 25, 2025 to January 5, 2026 were collected using Tweet-Harvest and preprocessed through *case folding*, *cleaning*, and *normalization*. Labeling was conducted using a combined approach of *pseudo-labeling* with a *confidence threshold* of 0.7 and manual annotation, with inter-annotator agreement measured by Cohen's Kappa of 0.808. The IndoBERT model was *fine-tuned* on 70% training data and evaluated on 884 manually verified test samples as *gold standard*. Results show that negative sentiment dominates at 50.4%, followed by neutral at 34.6% and positive at 15.0%. The model achieved *accuracy*, *macro F1-score*, and *weighted F1-score* of 0.91, demonstrating strong classification performance. These findings can help map public responses as input for education policy evaluation and communication. This study recommends clearer, phased, and more transparent socialization of the TKA policy.

Keywords: academic ability test, *fine-tuning*, IndoBERT, sentiment analysis, X platform

1. PENDAHULUAN

Evaluasi akademik merupakan instrumen penting dalam sistem pendidikan untuk menilai capaian pembelajaran peserta didik dan mendukung peningkatan mutu pendidikan secara berkelanjutan [1]. Kebijakan evaluasi pendidikan di Indonesia terus mengalami pembaruan, yang terbaru diwujudkan melalui penggantian Ujian Nasional (UN) menjadi Tes Kemampuan Akademik (TKA) sebagaimana diatur dalam Peraturan Menteri Pendidikan Dasar dan Menengah Nomor 9 Tahun 2025 [2]. Pada jenjang SMA, TKA memiliki implikasi signifikan karena berkaitan langsung dengan kesiapan akademik siswa dalam menentukan arah pendidikan lanjutan, sehingga

kebijakan ini tidak hanya berdampak pada peserta didik, tetapi juga memengaruhi orang tua, pendidik, dan masyarakat luas.

Sebagai kebijakan publik di bidang pendidikan, implementasi TKA memunculkan beragam respons dari berbagai pemangku kepentingan. Sebagian pihak mendukung TKA sebagai langkah maju dalam standarisasi penilaian akademik nasional, sementara sebagian lainnya menyampaikan kekhawatiran terkait kesiapan infrastruktur, beban psikologis siswa, ketidakselarasan kurikulum dengan materi yang diujikan, serta kejelasan fungsi hasil TKA dalam proses seleksi perguruan tinggi [3], [4]. Dinamika respons publik yang beragam ini menunjukkan perlunya pemetaan sistematis terhadap kecenderungan opini masyarakat sebagai dasar evaluasi dan komunikasi kebijakan.

Media sosial, khususnya platform X (sebelumnya Twitter), menjadi ruang utama bagi masyarakat untuk menyampaikan opini secara terbuka dan *real-time*. Berdasarkan laporan DataReportal, jumlah pengguna platform X di Indonesia mencapai 25,2 juta pada awal tahun 2025, dengan 19% di antaranya berusia 18–24 tahun yang beririsan dengan kelompok peserta didik SMA [5]. Opini yang disampaikan melalui platform X bersifat spontan dan merefleksikan persepsi masyarakat terhadap isu sosial, politik maupun kebijakan yang sedang berlangsung, sehingga platform ini menjadi sumber data yang potensial untuk mengkaji respons publik terhadap kebijakan TKA [6][7].

Meskipun diskursus mengenai TKA telah berkembang luas di ruang digital, opini publik tersebut sejauh ini belum dianalisis secara sistematis dengan pendekatan kuantitatif. Kondisi ini menyulitkan pembuat kebijakan dalam merancang respons, komunikasi risiko, dan strategi penyesuaian kebijakan yang responsif terhadap persepsi publik [8]. Analisis sentimen sebagai salah satu teknik dalam *Natural Language Processing* (NLP) menawarkan pendekatan komputasional untuk mengklasifikasikan opini ke dalam kategori positif, negatif, atau netral secara otomatis, sehingga pemetaan kecenderungan sentimen publik dapat dilakukan secara sistematis dan terukur [9].

Namun, data teks dari platform X memiliki karakteristik yang menantang secara teknis, yaitu bersifat pendek, informal, mengandung singkatan, bahasa gaul, serta ekspresi emosional yang implisit. Pada konteks kebijakan TKA, kompleksitas ini semakin meningkat karena munculnya istilah teknis pendidikan yang dipadukan dengan bahasa sehari-hari remaja dalam satu kalimat. Karakteristik ini menjadi tantangan bagi metode analisis sentimen tradisional berbasis leksikon yang bergantung pada fitur kata sederhana dan kurang mampu menangkap makna kontekstual [10].

Untuk mengatasi tantangan tersebut, penelitian ini menggunakan IndoBERT, yaitu model bahasa pralatih berbasis arsitektur *transformer* yang dikembangkan khusus untuk bahasa Indonesia [11]. IndoBERT memiliki kemampuan memahami konteks kata secara dua arah (*bidirectional*) sehingga mampu menangani variasi bahasa informal, ambiguitas semantik, dan ekspresi kontekstual yang umum ditemukan pada teks media sosial [12]. Sejumlah penelitian menunjukkan efektivitas IndoBERT dalam analisis sentimen pada data berbahasa Indonesia, khususnya pada data media sosial yang bersifat pendek dan informal [13], [14].

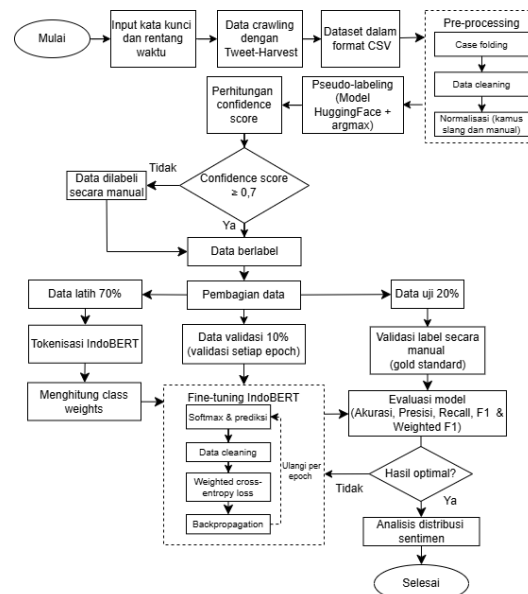
Berbagai kajian terdahulu telah menerapkan analisis sentimen terhadap kebijakan publik menggunakan data media sosial, namun terdapat kesenjangan penelitian yang signifikan. Penelitian yang mengkaji sentimen terhadap kebijakan pendidikan seperti MBKM masih menggunakan model LSTM yang memiliki keterbatasan dalam menangkap konteks semantik [15], sementara penelitian berbasis IndoBERT pada kebijakan publik seperti kenaikan PPN tidak spesifik pada domain pendidikan [13]. Hingga saat ini, belum terdapat penelitian yang secara khusus menganalisis sentimen terhadap kebijakan TKA jenjang SMA menggunakan IndoBERT pada data platform X.

Berdasarkan kesenjangan tersebut, penelitian ini bertujuan untuk: (1) menggambarkan distribusi sentimen pengguna media sosial X terhadap kebijakan TKA jenjang SMA berdasarkan klasifikasi sentimen positif, negatif, dan netral menggunakan model IndoBERT varian *indobert-base-p2*; serta (2) menganalisis kinerja model IndoBERT dalam mengklasifikasikan sentimen berdasarkan metrik *accuracy*, *precision*, *recall*, *F1-score*, dan *weighted F1-score*. Hasil penelitian diharapkan dapat menjadi masukan berbasis data bagi pemangku kepentingan dalam mengevaluasi dan merumuskan strategi komunikasi kebijakan TKA yang lebih responsif terhadap persepsi masyarakat.

2. METODE PENELITIAN

Data yang digunakan dalam penelitian ini berasal dari platform X berupa cuitan berbahasa Indonesia yang membahas kebijakan Tes Kemampuan Akademik (TKA) jenjang SMA. Pengumpulan data dilakukan menggunakan teknik *data crawling* dengan pustaka Tweet-Harvest berdasarkan kata kunci seperti "TKA SMA", "Tes Kemampuan Akademik", "simulasi TKA", dan "ujian TKA" pada periode 25 Mei 2025 hingga 5 Januari 2026. Label data yang digunakan dalam penelitian ini adalah positif, netral, dan negatif, yang diperoleh melalui pendekatan gabungan *pseudo-labeling* dengan *confidence threshold* 0,7 dan anotasi manual untuk data di bawah ambang batas tersebut.

Penelitian ini dilaksanakan melalui beberapa tahapan sistematis untuk memetakan sentimen pengguna media sosial X terhadap kebijakan TKA jenjang SMA berdasarkan klasifikasi sentimen positif, negatif, dan netral. Setiap langkah dalam proses ini digambarkan melalui sebuah diagram alur, seperti yang diperlihatkan pada Gambar 1 berikut:



Gambar 1. Diagram Alur Penelitian

Adapun tahapan metode dalam penelitian ini meliputi:

1. Pre-processing Teks

Pre-processing dilakukan untuk meningkatkan kualitas data sebelum dianalisis. Tahapan yang diterapkan meliputi *case folding* (mengubah seluruh teks menjadi huruf kecil), *data cleaning* (menghapus URL, *mention*, *hashtag*, emoji, angka, dan simbol), serta normalisasi menggunakan kamus *Kamus_Slang* dari Kaggle yang diperkaya secara manual dengan entri spesifik untuk konteks TKA. Tahap *stemming* dan penghapusan *stopword* tidak dilakukan untuk mempertahankan konteks semantik teks.

2. Pelabelan Data

Pelabelan dilakukan menggunakan pendekatan *pseudo-labeling* dengan model *indonesia-bert-sentiment-classification* melalui HuggingFace dengan *confidence threshold* 0,7. Data dengan *confidence score* di bawah threshold dilabeli secara manual. Tingkat kesesuaian antara label otomatis dan manual diukur menggunakan Cohen's Kappa dan memperoleh nilai 0,808 (kategori kuat).

3. Pembagian Dataset

Dataset dibagi dengan rasio 70% data latih (3.094 data), 10% data validasi (442 data), dan 20% data uji (884 data). Data uji diverifikasi seluruhnya melalui anotasi manual sebagai *gold standard* untuk memastikan validitas evaluasi model.

4. Fine-tuning IndoBERT

Fine-tuning dilakukan menggunakan model *indobenchmark/indobert-base-p2* dengan tiga kelas keluaran. Teks ditokenisasi menggunakan *WordPiece tokenizer* dengan penambahan token [CLS] sebagai representasi keseluruhan kalimat. Vektor [CLS] diproses melalui lapisan klasifikasi linear dan fungsi *softmax* untuk menghasilkan probabilitas per kelas. Pelatihan menggunakan optimizer AdamW dengan *learning rate* 2×10^{-5} , *batch size* 16, *weight decay* 0,05, dan 4 *epoch*. Mengingat distribusi kelas yang tidak seimbang, diterapkan *weighted cross-entropy loss* dengan bobot kelas dihitung menggunakan *inverse class frequency*.

5. Evaluasi Model

Evaluasi dilakukan menggunakan 884 data uji berlabel *gold standard* berdasarkan *confusion matrix* dan metrik *accuracy*, *precision*, *recall*, *macro F1-score*, dan *weighted F1-score*. Seluruh implementasi menggunakan Python 3.12.13 pada Google Colaboratory dengan GPU T4.

3. HASIL DAN PEMBAHASAN

3.1. Hasil

Proses *data crawling* menghasilkan 4.439 cuitan yang setelah *pre-processing* diperoleh 4.420 data bersih. Tahap *pre-processing* meliputi *case folding*, *data cleaning*, dan normalisasi menggunakan kamus *Kamus Slang* dari Kaggle yang diperkaya secara manual dengan entri spesifik konteks TKA. Seluruh tahapan ini dilakukan tanpa *stemming* dan penghapusan *stopword* untuk mempertahankan konteks semantik yang dibutuhkan model IndoBERT. Data yang sudah melalui *pre-processing* selanjutnya masuk ke tahap pelabelan data.

Pelabelan data menghasilkan distribusi sentimen yang tidak seimbang sebagaimana ditampilkan pada Tabel 1. Sentimen negatif mendominasi dataset dengan 2.229 data (50,4%), diikuti netral 1.528 data (34,6%), dan positif 663 data (15,0%).

Tabel 1. Distribusi Label Sentimen Dataset Akhir

Label Sentimen	Jumlah Data	Persentase
Negatif	2.229	50,4%
Netral	1.528	34,6%
Positif	663	15,0%
Total	4.420	100%

Untuk memperkuat gambaran distribusi sentimen, dilakukan visualisasi kata dalam bentuk word cloud pada masing-masing kategori sebagaimana ditampilkan pada Gambar 2.



Gambar 2. Visualisasi Word Cloud pada Setiap Kategori Sentimen

Berdasarkan visualisasi tersebut, sentimen negatif didominasi kata-kata seperti "susah", "bingung", "takut", dan "curang"; sentimen netral didominasi kata-kata informatif seperti "pelajaran", "materi", dan "jurusan"; sedangkan sentimen positif didominasi ekspresi harapan seperti "semoga", "alhamdulillah", dan "semangat".

Dari 4.420 data, sebanyak 3.769 data memperoleh *pseudo-label* otomatis dengan *confidence score* $\geq 0,7$, sedangkan 651 data sisanya dilabeli secara manual. Tingkat kesesuaian antara *pseudo-label* dan label manual pada data uji diukur menggunakan Cohen's Kappa dan memperoleh nilai 0,808, yang termasuk dalam kategori kuat. Dataset kemudian dibagi dengan rasio 70:10:20 menjadi 3.094 data latih, 442 data validasi, dan 884 data uji. Seluruh 884 data uji diverifikasi melalui anotasi manual sebagai *gold standard* dengan distribusi sebagaimana ditampilkan pada Tabel 2.

Tabel 2. Distribusi Label Gold Standard pada Data Uji

Label Sentimen	Jumlah Data	Persentase
----------------	-------------	------------

Positif	147	16,6%
Netral	309	35,0%
Negatif	428	48,4%
Total	884	100%

Proses *fine-tuning* IndoBERT dilakukan selama 4 *epoch* dengan konfigurasi *learning rate* 2×10^{-5} , *batch size* 16, dan *weight decay* 0,05. Perkembangan nilai *training loss*, *validation loss*, dan *weighted F1-score* pada setiap *epoch* ditampilkan pada Tabel 3.

Tabel 3. Hasil Pelatihan IndoBERT per Epoch

Epoch	Training Loss	Validation Loss	Accuracy	Weighted F1
1	0,346787	0,376745	0,848416	0,852272
2	0,282170	0,425926	0,884615	0,884472
3	0,106591	0,577480	0,884615	0,884577
4	0,013568	0,576969	0,880090	0,880162

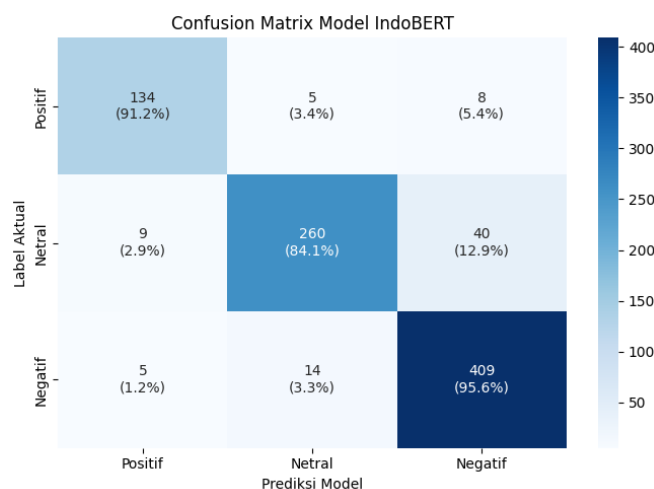
Model terbaik dipilih berdasarkan *weighted F1-score* tertinggi pada data validasi, yang dicapai pada epoch ke-3 dengan nilai 0,8846. Terdapat indikasi *overfitting* mulai epoch ke-3 yang ditandai oleh penurunan *training loss* secara signifikan sementara *validation loss* terus meningkat. Kondisi ini diantisipasi melalui penerapan *weight decay* 0,05, *dropout* 0,1, dan mekanisme *load_best_model_at_end* yang secara otomatis memuat model terbaik berdasarkan *weighted F1-score* pada data validasi.

Evaluasi akhir dilakukan menggunakan 884 data uji berlabel *gold standard*. Hasil evaluasi per kelas ditampilkan pada Tabel 4.

Tabel 4. Hasil Evaluasi Model IndoBERT pada Data Uji

Kelas	Precision	Recall	F1-score	Support
Positif	0,91	0,91	0,91	147
Netral	0,93	0,84	0,88	309
Negatif	0,89	0,96	0,92	428
Accuracy			0,91	884
Macro Avg	0,91	0,90	0,91	884
Weighted Avg	0,91	0,91	0,91	884

Model memperoleh *accuracy* 0,91, *macro F1-score* 0,91, dan *weighted F1-score* 0,91. Hasil *confusion matrix* ditampilkan pada Gambar 3.



Gambar 3. Confusion Matrix Model IndoBERT

Dari 884 data uji, sebanyak 803 data berhasil diklasifikasikan dengan benar. Kesalahan terbanyak terjadi pada data aktual netral yang diprediksi sebagai negatif, yaitu sebanyak 40 data (12,9% dari total data aktual netral).

3.2. Pembahasan

Dominasi sentimen negatif sebesar 50,4% dalam dataset penelitian ini mencerminkan kecenderungan respons publik yang lebih banyak bersifat kritis terhadap kebijakan TKA jenjang SMA. Temuan ini konsisten dengan pola yang ditemukan pada penelitian-penelitian sebelumnya mengenai kebijakan pendidikan, di mana kebijakan baru yang memiliki implikasi langsung terhadap peserta didik cenderung memunculkan resistensi dan kekhawatiran di ruang publik [15]. Berdasarkan visualisasi *word cloud*, sentimen negatif didominasi oleh kata-kata seperti "susah", "bingung", "takut", "stres", "pasrah", dan "curang", yang mengindikasikan bahwa respons negatif tidak hanya berkaitan dengan penolakan terhadap kebijakan, tetapi juga dengan tekanan psikologis, ketidakpastian teknis pelaksanaan, serta kekhawatiran terhadap integritas pelaksanaan TKA. Temuan ini sejalan dengan penelitian [3] yang mengidentifikasi bahwa implementasi TKA pada tahap awal menghadapi berbagai tantangan terkait kesiapan infrastruktur dan dukungan pemangku kepentingan. Sementara itu, sentimen positif yang muncul dalam bentuk harapan dan ungkapan syukur menunjukkan bahwa sebagian masyarakat tetap merespons kebijakan ini secara optimistis, meskipun proporsinya jauh lebih kecil dibandingkan sentimen negatif.

Keunggulan *contextual embedding* IndoBERT terbukti relevan dalam konteks data media sosial berbahasa Indonesia yang bersifat informal dan dinamis. Berbeda dengan pendekatan *static embedding* seperti Word2Vec yang menghasilkan satu representasi vektor tetap untuk setiap kata, IndoBERT menghasilkan representasi yang berubah sesuai konteks kalimat (Wilie et al., 2020). Hal ini memungkinkan model membedakan makna kata yang sama dalam konteks berbeda, misalnya kata "belajar" yang dapat bersifat netral dalam konteks informasi, tetapi bernuansa negatif ketika muncul bersama ekspresi seperti "susah" atau "takut". Kemampuan ini menjadi keunggulan IndoBERT dibandingkan metode berbasis leksikon yang hanya mengandalkan polaritas kata secara individual tanpa mempertimbangkan konteks kalimat.

Model IndoBERT varian *indobert-base-p2* menghasilkan *accuracy* 0,91 dan *weighted F1-score* 0,91 pada data uji berlabel *gold standard*, menunjukkan bahwa model mampu mengklasifikasikan sentimen pada teks informal berbahasa Indonesia dengan performa yang baik. Hasil ini memberikan bukti empiris bahwa pendekatan *fine-tuning* model *transformer* berbasis bahasa Indonesia efektif diterapkan pada data media sosial dalam konteks kebijakan pendidikan, melengkapi penelitian sebelumnya yang menunjukkan efektivitas serupa pada domain lain seperti kebijakan fiskal [13] dan isu pemerintahan [16]. Keunggulan IndoBERT dalam konteks ini terletak pada kemampuan *contextual embedding* yang menghasilkan representasi vektor kata secara dinamis sesuai konteks kalimat, berbeda dengan pendekatan *static embedding* seperti Word2Vec yang menghasilkan representasi tetap tanpa mempertimbangkan konteks [12]. Kemampuan ini memungkinkan model membedakan makna yang berbeda dari kata yang sama dalam konteks berbeda, misalnya kata "susah" yang dapat bersifat informatif dalam satu kalimat tetapi bernuansa keluhan dalam kalimat lain. Namun demikian, kemampuan memahami konteks semantik belum sepenuhnya mencakup konteks pragmatik. Kesalahan terbanyak terjadi pada cuitan netral yang diprediksi sebagai negatif (40 data; 12,9%), yang menunjukkan bahwa model masih menghadapi tantangan dalam membedakan cuitan informatif bernuansa kecemasan dari keluhan eksplisit yang merupakan sebuah keterbatasan yang juga diidentifikasi pada model berbasis BERT secara umum dalam memahami ekspresi implisit dan ironi [17].

Penerapan *weighted cross-entropy loss* dengan bobot kelas berbasis *inverse class frequency* dalam penelitian ini merupakan respons terhadap ketidakseimbangan distribusi kelas yang signifikan. Strategi ini mengacu pada prinsip bahwa model yang dilatih pada data tidak seimbang cenderung bias terhadap kelas mayoritas. Hasil evaluasi menunjukkan bahwa kelas positif sebagai kelas minoritas memperoleh F1-score 0,91, yang sebanding dengan kelas negatif (0,92) dan lebih tinggi dari kelas netral (0,88). Hal ini mengindikasikan bahwa pembobotan kelas berkontribusi dalam menjaga keseimbangan performa antar kelas, meskipun efektivitasnya secara komparatif masih memerlukan pengujian lanjutan dengan membandingkan hasil pelatihan dengan dan tanpa pembobotan.

Temuan penelitian ini memiliki implikasi praktis yang relevan bagi pemangku kebijakan pendidikan. Dominasi sentimen negatif yang diiringi kemunculan kata "curang" dan "bingung" mengisyaratkan dua kebutuhan mendesak: pertama, penguatan strategi komunikasi dan sosialisasi kebijakan TKA yang lebih terstruktur, bertahap, dan mudah dipahami oleh siswa dan orang tua; kedua, evaluasi terhadap sistem integritas pelaksanaan TKA yang mencakup prosedur pengawasan, validasi peserta, dan transparansi penilaian. Secara lebih luas, penelitian ini mendemonstrasikan bahwa analisis sentimen berbasis *transformer* dapat menjadi instrumen komunikasi risiko yang efektif bagi pembuat kebijakan dalam membaca dinamika opini publik secara kuantitatif dan berbasis data, sebagaimana direkomendasikan oleh [8] dalam perumusan kebijakan yang responsif terhadap persepsi masyarakat.

4. KESIMPULAN

Berdasarkan hasil penelitian dan pembahasan, analisis sentimen berbasis IndoBERT dapat diterapkan secara efektif untuk memetakan kecenderungan opini publik terhadap kebijakan pendidikan dari data media sosial

berbahasa Indonesia yang bersifat informal dan tidak terstruktur. Dominasi sentimen negatif sebesar 50,4% dari total 4.420 cuitan terkait kebijakan TKA jenjang SMA mengindikasikan bahwa respons masyarakat lebih banyak diwarnai oleh kekhawatiran, ketidakpastian teknis, tekanan akademik, dan perhatian terhadap integritas pelaksanaan. Temuan ini menegaskan urgensi pemetaan opini publik secara sistematis sebagai instrumen komunikasi risiko bagi pemangku kebijakan, terutama pada fase awal implementasi kebijakan evaluasi pendidikan yang bersifat nasional.

Model IndoBERT varian *indobert-base-p2* menghasilkan *weighted F1-score* 0,91 pada data uji berlabel *gold standard*, menunjukkan bahwa pendekatan *transfer learning* berbasis *transformer* mampu menangani kompleksitas bahasa informal media sosial berbahasa Indonesia dengan baik. Keterbatasan utama terletak pada perbedaan cuitan netral dari negatif dalam konteks pragmatik, khususnya pada cuitan yang mengandung ekspresi kecemasan namun secara keseluruhan bersifat informatif.

DAFTAR PUSTAKA

- [1] T. Akbar *et al.*, “Analisis Kritis Sistem Evaluasi Pendidikan Indonesia,” *Ma’arif J. Educ.*, vol. 1, no. 1, pp. 1–11, 2022.
- [2] *Peraturan Menteri Pendidikan Dasar dan Menengah Republik Indonesia Nomor 9 Tahun 2025 tentang Tes Kemampuan Akademik*. Kementerian Pendidikan Dasar dan Menengah Republik Indonesia, 2025.
- [3] Y. Hertanto and H. Prasetyono, “Paradoks Standardisasi Akademik dalam Pendidikan Vokasi : Analisis Kritis Kebijakan TKA Melalui Studi Dokumen,” *JPI J. Pendidik. Indones.*, vol. 5, no. 6, 2025, doi: 10.59818/jpi.v5i6.2320.
- [4] J. Juriyah, T. Fahrudin, and J. Jalaludin, “Readiness Analysis of the Implementation of the 2025 Academic Ability Test (TKA) Policy by the Ministry of Education and Culture,” *YASIN J. Pendidik. dan Sos. Budaya*, vol. 5, pp. 5708–5721, 2025.
- [5] “Digital 2025: Indonesia — DataReportal.” Accessed: Sep. 02, 2026. [Online]. Available: <https://datareportal.com/reports/digital-2025-indonesia>
- [6] D. E. Ratnawati, N. Y. Setiawan, P. Studi, S. Informasi, F. I. Komputer, and U. Brawijaya, “Analisis Sentimen Pengguna Sosial Media Twitter / X Terhadap Acara Clash of Champions Menggunakan,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 9, no. 3, pp. 1–10, 2025.
- [7] F. Z. Emeraldien, R. J. Sunarsono, and R. Alit, “TWITTER SEBAGAI PLATFORM KOMUNIKASI POLITIK DI INDONESIA,” *SCAN*, vol. 14, no. 1, 2019.
- [8] R. A. Faridz, “ANALISIS SENTIMEN TWITTER MENGGUNAKAN ALGORITMA K-MEANS MENGENAI DAMPAK PENGHAPUSAN TKA SNBT TAHUN 2023,” 2024.
- [9] H. I. Ariq and S. L. Siregar, “Analisis Sentimen Ulasan Produk untuk Mengukur Kepuasan Pelanggan E-commerce menggunakan metode Natural Language Processing (NLP) dan Support Vector Machine (SVM),” *JPNM (Jurnal Pustaka Nusantara. Multidisiplin)*, vol. 3, no. 3, pp. 1–8, 2025.
- [10] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding Jacob,” in *Proceedings of the 17th Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2019, p. 16. doi: 10.18653/v1/N19-1423.
- [11] A. Sevtian, Y. B. Pratama, and Z. Wahyuzi, “Analisis Sentimen terhadap Kasus Korupsi Timah di Kepulauan Bangka Belitung menggunakan Algoritma Indobert dan Bidirectional LSTM,” *Hum. J. Soc. Humanit. Sci.*, vol. 3, no. 1, pp. 42–53, 2025, doi: <https://doi.org/10.58738/human.v3i1.1116>.
- [12] 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*, 2020, pp. 843–857.
- [13] M. R. Manoppo, I. C. Kolang, D. N. Fiat, R. Michelly, and C. Mawara, “ANALISIS SENTIMEN PUBLIK DI MEDIA SOSIAL TERHADAP KENAIKAN PPN 12 % DI INDONESIA MENGGUNAKAN INDOBERT ANALYSIS OF PUBLIC SENTIMENT ON SOCIAL MEDIA TOWARDS THE 12 % PPN INCREASE IN INDONESIA USING INDOBERT,” *J. Kecerdasan Buatan dan Teknol. Inf.*, vol. 4, no. 2, pp. 152–163, 2025, doi: <https://doi.org/10.69916/jkbt.v4i2.322>.
- [14] T. Marwiah, D. A. Ansani, M. Fitriani, and A. Laia, “Analisis Sentimen Komentar YouTube pada Video Jawaban Kobo Purbaya Menggunakan IndoBERT,” *Metri. J.*, vol. 6, no. 1, 2025.

-
- [15] sio J. Pipin and H. Kurniawan, “Analisis Sentimen Kebijakan MBKM Berdasarkan Opini Masyarakat di Twitter Menggunakan LSTM,” *J. SIFO Mikroskil*, vol. 23, no. 2, pp. 197–208, 2022, doi: 10.55601/jsm.v23i2.900.
- [16] S. I. Ishak, O. Arnilia, T. Widodo, I. G. Nyoman, and A. Bisma, “Analisis sentimen terhadap pemerintahan Prabowo – Gibran menggunakan IndoBERT dan LDA,” *JAMBURA J. INFORMATICS*, vol. 7, no. 2, pp. 72–82, 2025, doi: 10.37905/jji.v1i2.34895.
- [17] P. D. Manalu, M. Simanjuntak, and C. Umri, “Implementasi Algoritma Klasifikasi untuk Analisis Sentimen Media Sosial Tiktok Tahun 2025 analisis,” *J. Tek. Inform. dan Teknol. Inf.*, vol. 5, no. 1, 2025, doi: <https://doi.org/10.55606/jutiti.v5i1.5644>.