Digital Transformation Technology (Digitech) | e-ISSN: 2807-9000
Volume 6, **Number** 1, Maret 2026
https://doi.org/10.47709/digitech.v6i1.7909

Klasifikasi Sentimen Publik di Instagram Menggunakan Indobert Terhadap Respons DPR Pasca-Aksi '17+8 Tuntutan Rakyat

Haddad Alwi[1], M.Julkarnain[2*], Siska Atmawan Oktavia[3], Dinola[4]
[1234]Universitas Teknologi Sumbawa, Indonesia
[1]haddadalwi864@email.com, [2]m.julkarnain@uts.ac.id, [3]siska.atmawan.oktavia@uts.a.id [4]dinola@uts.ac.id

open access
Histori Artikel:
Diajukan: 27 Januari 2026
Disetujui: 15 Februari 2026
Dipublikasi: 20 Februari 2026
Kata Kunci:
Analisis Sentimen, IndoBERT, Instagram, DPR RI, KDD
Digital Transformation Technology (Digitech) *is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).*

Abstrak
Penelitian ini bertujuan untuk mengevaluasi kinerja model IndoBERT dalam klasifikasi sentimen tiga kelas (positif, negatif, dan netral) serta memetakan persepsi publik terhadap respons DPR RI pasca aksi "17+8 Tuntutan Rakyat" berdasarkan komentar berbahasa Indonesia di platform Instagram. Pendekatan Knowledge Discovery in Databases (KDD) digunakan yang meliputi tahapan seleksi data (4.737 komentar Instagram), praproses teks (cleansing, case folding, dan normalisasi bahasa gaul), pelabelan data secara semi-otomatis dengan metode Lexicon-Based, pemodelan, dan evaluasi. Untuk mengatasi ketidakseimbangan data (class imbalance), penelitian ini menerapkan teknik class weighting. Model IndoBERT (indobert-base-p1) diterapkan melalui proses fine-tuning dengan rasio pembagian dataset 70:15:15, menggunakan optimizer AdamW, learning rate 2e-5, dan 4 epochs. Evaluasi kinerja model dilakukan menggunakan metrik akurasi, presisi, recall, dan F1-score melalui confusion matrix. Hasil penelitian menunjukkan bahwa IndoBERT mencapai akurasi sebesar 97,47% dengan nilai presisi, recall, dan F1-score yang seimbang pada setiap kelas sentimen. Pemetaan sentimen mengindikasikan bahwa sentimen netral mendominasi respons publik, diikuti oleh sentimen negatif dan positif. Selain itu, analisis word cloud menunjukkan perbedaan karakteristik linguistik pada masing-masing kategori sentimen. Temuan ini menegaskan bahwa IndoBERT memiliki kinerja yang sangat baik dalam mengklasifikasikan dan memetakan sentimen publik terhadap isu sosial-politik pada media sosial berbahasa Indonesia.

PENDAHULUAN

Perkembangan teknologi informasi dan komunikasi telah mendorong media sosial menjadi ruang publik digital yang berperan penting dalam pembentukan opini dan diskursus politik masyarakat. Instagram, sebagai platform dengan tingkat interaksi visual dan tekstual yang tinggi di Indonesia, menjadi kanal utama bagi pengguna untuk merespons kebijakan lembaga negara (Halberstadt et al., 2022). Fenomena ini terlihat pasca-aksi "17+8 Tuntutan Rakyat", di mana pernyataan Dewan Perwakilan Rakyat (DPR) RI memicu gelombang komentar masif yang mencerminkan berbagai spektrum emosi, mulai dari dukungan hingga kritik tajam (Wikipedia, 2025).

Analisis sentimen terhadap isu politik di Indonesia telah banyak dilakukan untuk membedah opini publik secara otomatis. Penelitian terdahulu oleh(Khairani et al., 2024) menunjukkan bahwa penggunaan algoritma machine learning tradisional dapat memberikan gambaran sentimen yang cukup baik, namun sering kali terkendala oleh ambiguitas bahasa. Lebih lanjut,(Geni et al., 2023) membuktikan bahwa model berbasis Transformer seperti IndoBERT memiliki keunggulan signifikan dalam menangkap konteks kalimat dibandingkan metode konvensional pada data media sosial Twitter. Meskipun demikian, sebagian besar studi literatur masih berfokus pada Twitter (X), sementara analisis pada platform Instagram yang memiliki karakteristik bahasa lebih santai, penuh singkatan, dan penggunaan dialek lokal yang lebih beragam, masih relatif terbatas (Setiawan, 2025).

Tantangan utama dalam menganalisis komentar Instagram adalah sifat datanya yang sangat dinamis dan tidak terstruktur. Penggunaan singkatan dan bahasa gaul sering kali menurunkan performa metode klasifikasi tradisional (Koto, 2021).Untuk mengatasi celah tersebut, penelitian ini mengadopsi model *Indonesian Bidirectional Encoder Representations from Transformers* (IndoBERT). Berbeda dengan model pendahulunya, IndoBERT dilatih pada korpus bahasa Indonesia yang sangat besar dan menggunakan mekanisme *self-attention* dua arah, sehingga mampu memahami makna kata berdasarkan konteks kalimat secara lebih mendalam (Devlin et al., 2018).

Penelitian ini memfokuskan pada pengujian efektivitas model representasi bahasa pra-latih tersebut dalam memetakan dinamika opini publik yang kompleks secara lebih akurat dan objektif. Melalui pendekatan ini, diharapkan diperoleh gambaran berbasis data mengenai persepsi masyarakat terhadap respons lembaga legislatif, sekaligus memberikan kontribusi pada pengembangan metodologi klasifikasi teks otomatis untuk bahasa

Indonesia di media sosial yang memiliki tantangan linguistik tinggi.

STUDI LITERATUR

Penelitian Terkait

Berbagai penelitian telah mengkaji analisis sentimen pada media sosial menggunakan model berbasis Transformer. (Khairani et al., 2024) menemukan bahwa IndoBERT memiliki kemampuan kuat dalam memahami konteks media sosial tanpa praproses yang kompleks, mencapai akurasi 92,54%. Sementara itu, (Geni et al., 2023) membuktikan efektivitas IndoBERT dalam menangani ambiguitas bahasa politik di Twitter dengan akurasi 89%. Berbeda dengan penelitian sebelumnya yang mayoritas berfokus pada Twitter, penelitian ini mengisi celah dengan menganalisis platform Instagram yang memiliki karakteristik linguistik lebih dinamis untuk memetakan sentimen publik terhadap lembaga legislatif.

Instagram

Instagram telah bertransformasi menjadi ruang diskursus digital yang merefleksikan opini masyarakat terhadap isu sosial dan politik secara langsung melalui kolom komentarnya (Prajarini, 2021). Karakteristik komentarnya yang singkat, informal, dan kaya akan ekspresi nonbaku menjadikannya data yang sangat relevan untuk memetakan sentimen publik secara *real-time*. Untuk mengumpulkan data masif tersebut secara efisien dan menghindari bias manual, digunakan teknik *web scraping* (Sheikh et al., 2024). Pendekatan ini memungkinkan pengambilan ribuan opini publik terkait aksi "17+8 Tuntutan Rakyat" secara otomatis untuk kemudian digunakan sebagai input utama dalam tahapan analisis

Metodologi Knowledge Discovery in Databases (KDD)

Penelitian ini mengadopsi kerangka kerja KDD sebagai panduan prosedural untuk mengolah data tekstual secara sistematis. Tahapan yang dilakukan meliputi seleksi data (pengumpulan komentar via *web scraping*), praproses (pembersihan dan normalisasi), transformasi, hingga *data mining* menggunakan model IndoBERT (Suparwito et al., 2024). Penggunaan KDD memastikan ekstraksi wawasan dari komentar Instagram dilakukan secara terstruktur guna menghasilkan data yang akurat (Sari et al., 2024).

Pra-pemrosesan

Tahap pra-pemrosesan merupakan langkah krusial untuk mentransformasi komentar Instagram yang tidak terstruktur menjadi format yang optimal bagi model pembelajaran mesin (Kumar et al., 2025). Proses ini mencakup teknik *cleaning* untuk menghapus atribut non-teks, *case folding*, serta normalisasi untuk memperbaiki kesalahan pengetikan menjadi bentuk standar bahasa Indonesia (Yessy Asri & Dr. Dra. Dwina Kuswardani, 2024). Rangkaian prosedur ini berfungsi sebagai fondasi penting guna memastikan kualitas data referensi (*ground truth*) yang akan dipelajari oleh model, sehingga mengurangi *noise* yang dapat menurunkan akurasi prediksi sentimen.

IndoBERT

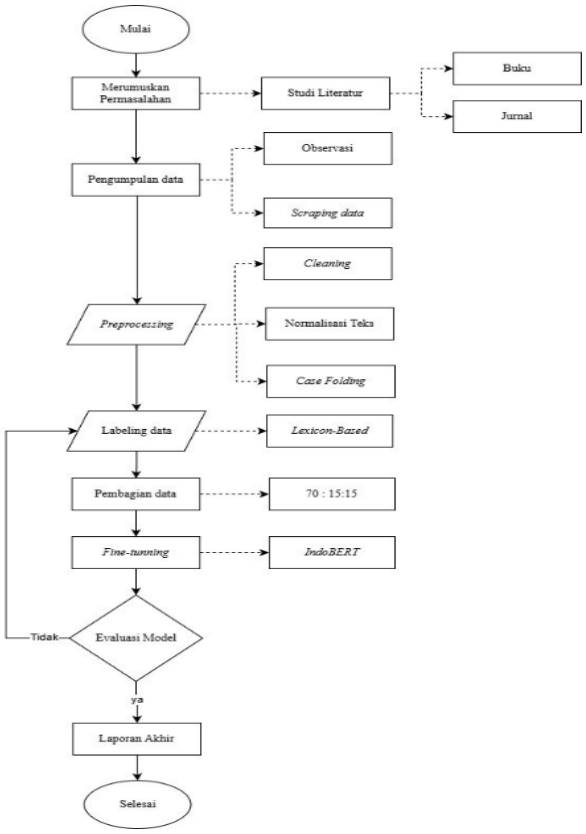
IndoBERT adalah model bahasa berbasis arsitektur Transformer yang dirancang khusus untuk menangkap nuansa linguistik bahasa Indonesia melalui mekanisme *self-attention* dan *bidirectional context* (Devlin et al., 2018). Model ini dipilih karena performanya yang unggul dalam menangani perpaduan bahasa formal dan non-formal yang mendominasi interaksi digital masyarakat Indonesia (Asri et al., 2025). Dengan kemampuan memahami hubungan antar-kata secara mendalam, IndoBERT digunakan dalam penelitian ini untuk mengklasifikasikan persepsi publik ke dalam kelas positif, negatif, dan netral dengan tingkat presisi yang lebih tinggi dibandingkan model konvensional.

Evaluasi Model

Untuk mengukur sejauh mana model IndoBERT mampu menggeneralisasi pola sentimen secara akurat, dilakukan tahap evaluasi menggunakan instrumen *Confusion Matrix*. Tabel evaluasi ini membandingkan hasil prediksi sistem dengan label asli pada data uji untuk menghasilkan metrik performa seperti akurasi, presisi, dan *recall* (Geron, 2022). Tahap ini menjadi sangat penting dalam konteks penelitian isu sosial-politik guna menjamin bahwa hasil pemetaan distribusi sentimen yang dihasilkan memiliki validitas ilmiah yang kuat sebelum diinterpretasikan lebih lanjut.

METODE

Penelitian ini menggunakan pendekatan Knowledge Discovery in Databases (KDD) sebagai metodologi sistematis dalam proses ekstraksi pengetahuan dari data teks tidak terstruktur. KDD dipilih karena menyediakan kerangka terintegrasi yang meliputi tahapan seleksi data, pra-pemrosesan, transformasi, penambangan data, serta evaluasi dan interpretasi hasil. Adapun tahapan penelitian disampaikan pada diagram alir pada Gambar 1. Diagram Alir



Gambar 1. Diagram Alir

Seleksi Data

Dataset yang digunakan dalam penelitian ini berupa komentar publik dari akun resmi DPR RI di Instagram pada unggahan yang membahas aksi "17+8 Tuntutan Rakyat." Data diperoleh melalui teknik web scraping menggunakan pustaka Python seperti BeautifulSoup dan Selenium. Proses ini menghasilkan 5.004 komentar yang disimpan dalam format Excel untuk memudahkan pengolahan selanjutnya.

Pra-pemrosesan

Tahap ini bertujuan untuk meningkatkan kualitas teks agar siap digunakan pada proses pemodelan dengan algoritma pembelajaran mesin. Proses pra-pemrosesan mencakup beberapa langkah penting, yaitu pembersihan data dari simbol, tag HTML, tautan, serta emoji yang tidak relevan (*cleaning*), penyeragaman seluruh huruf menjadi bentuk kecil (*case folding*), penghapusan kata umum yang tidak memiliki makna signifikan (*stopword removal*), dan normalisasi kata tidak baku agar sesuai dengan kaidah Bahasa Indonesia. Selain itu, dilakukan juga proses *labeling* atau pemberian kategori sentimen terhadap setiap data komentar berdasarkan konteks kalimatnya ke dalam tiga kelas, yaitu positif, negatif, dan netral. Hasil akhir dari tahap ini berupa dataset teks yang bersih, terstandarisasi, dan telah berlabel, sehingga siap untuk digunakan pada tahap transformasi selanjutnya.

Transformasi Data

Tahap transformasi dilakukan untuk mengubah teks menjadi format numerik agar dapat dibaca oleh model pembelajaran mesin. Proses *tokenization* dan *encoding* dilakukan menggunakan tokenizer bawaan IndoBERT, yang mengubah setiap kalimat menjadi representasi vektor berbasis *subword embeddings*. Dataset kemudian dibagi menjadi tiga subset, yaitu data latih (70%), validasi (15%), dan uji (15%).

Penambangan Data

Tahap inti penelitian dilakukan pada fase ini. Model *IndoBERT* digunakan untuk melakukan *fine-tuning* terhadap dataset hasil transformasi. Proses pelatihan dilakukan menggunakan *learning rate* sebesar 2e-5, *batch size* 16, dan jumlah *epoch* 4. Tujuan tahap ini adalah mengidentifikasi pola dan hubungan semantik antar kata sehingga model dapat melakukan klasifikasi sentimen secara otomatis.

Evaluasi dan Interpretasi

Tahapan akhir bertujuan untuk menilai performa model serta menginterpretasikan hasil klasifikasi. Evaluasi kuantitatif dilakukan menggunakan metrik akurasi, presisi, recall, dan F1-score untuk memastikan validitas model. Sedangkan tahap interpretasi dilakukan melalui analisis distribusi sentimen dan visualisasi word cloud pada setiap kelas sentimen. Hasil ini memberikan pemahaman mendalam mengenai kecenderungan opini publik terhadap respons DPR RI pasca aksi "17+8 Tuntutan Rakyat."

HASIL

Hasil Pengumpulan Data (Scraping)

Tahap awal penelitian dilakukan dengan mengekstraksi data dari platform Instagram guna menangkap opini publik secara langsung. Fokus pengumpulan data diarahkan pada unggahan yang berkaitan dengan respons DPR RI pasca aksi "17+8 Tuntutan Rakyat". Melalui teknik *web scraping* otomatis berbasis Python, penelitian ini berhasil menghimpun dataset yang cukup besar untuk menjamin representasi opini.

Proses tersebut menghasilkan total 5.004 komentar mentah. Setiap data yang ditarik mencakup atribut nama pengguna (*username*), isi komentar, dan tanggal unggah. Namun, untuk kepentingan analisis sentimen, atribut isi komentar dipilih sebagai sumber data utama karena mengandung ekspresi linguistik yang mencerminkan persepsi masyarakat.

Hasil Pra-processing Data

Mengingat karakteristik komentar Instagram yang cenderung kotor (mengandung emoji, bahasa gaul, dan simbol), tahap pra-pemrosesan menjadi krusial untuk meningkatkan kualitas data sebelum masuk ke tahap pemodelan. Langkah-langkah yang dilakukan meliputi pembersihan simbol, *case folding*, penghapusan *stopword*, hingga normalisasi kata tidak baku.
Perubahan signifikan pada data setelah melalui tahap ini dapat dilihat pada Tabel 1 di bawah ini:

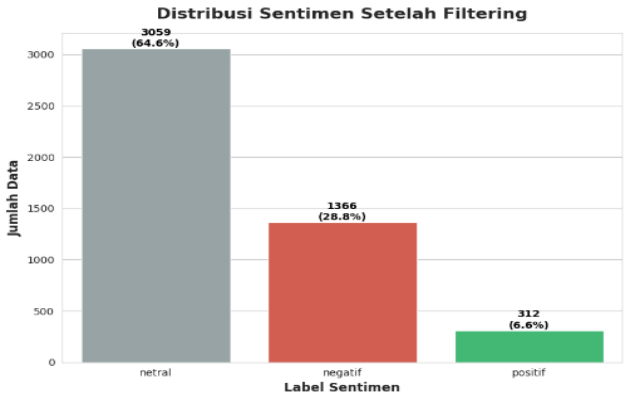
Tabel 1. Hasil Pra-processing

Komentar Asli	Komentar Hasil Pra-processing
KENAPA GA PUAN YANG MAJU,, HASILNYA JUGA GA JELAS GINI,, 😡😡😡 SIAP DIDEMO DENGAN AKSI YANG LEBIH DARI KEMAREN INI	kenapa tidak puan yang maju hasilnya juga tidak jelas gini siap didemo dengan aksi yang lebih dari kemarin ini
Wakil koq malah ngatur sendiri... 👏👏👏👏🐿️🐿️🐿️🐿️	wakil kok malah ngatur sendiri
Bapak ibuk PEJABAT kapan ruu perampasan aset disahkan ?	bapak ibuk pejabat kapan ruu perampasan aset disahkan

Berdasarkan Tabel 1, terlihat bahwa proses pra-pemrosesan berhasil mereduksi *noise*. Penghapusan emoji dan tanda baca membuat model dapat fokus pada konten tekstual, sementara normalisasi kata (seperti "ga" menjadi "tidak" atau "kemaren" menjadi "kemarin") membantu model IndoBERT mengenali kata dalam bentuk standarnya, sehingga meningkatkan akurasi pemahaman konteks.

Hasil Pelabelan Sentimen

Setelah data bersih, dilakukan proses anotasi atau pelabelan untuk menentukan kelas sentimen pada setiap komentar. Pelabelan dilakukan secara manual ke dalam tiga kategori: positif, negatif, dan netral, guna memastikan kualitas *ground truth* yang akan digunakan dalam pelatihan model.
Distribusi hasil pelabelan tersebut divisualisasikan dalam Gambar 2 berikut:



Gambar 2. Distribusi Labeling

Melalui visualisasi pada Gambar 2, terlihat bahwa kelas netral mendominasi dataset dibandingkan kelas positif dan negatif. Hal ini menunjukkan bahwa sebagian besar respons publik di kolom komentar DPR RI bersifat informatif, menanyakan kejelasan, atau memberikan pernyataan tanpa muatan emosional yang ekstrem. Dominasi

kelas netral ini menjadi tantangan tersendiri bagi model untuk tetap presisi dalam membedakan nuansa antara sentimen yang benar-benar netral dengan sentimen negatif yang bersifat satir.

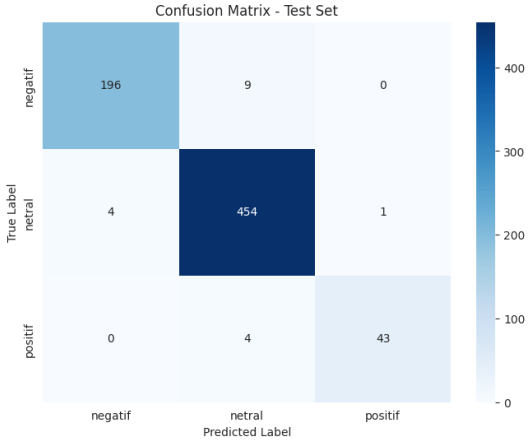
Hasil Pemodelan

Pada tahap pemodelan, penelitian ini menggunakan arsitektur IndoBERT yang telah melalui proses *fine-tuning*. Dataset dibagi dengan rasio 70% data latih, 15% data validasi, dan 15% data uji. Pembagian ini bertujuan agar model memiliki data yang cukup untuk belajar, namun tetap memiliki cadangan data yang memadai untuk menguji kemampuan generalisasinya terhadap data baru.

Proses *fine-tuning* dilakukan dengan menambahkan lapisan klasifikasi (*dense layer*) di atas arsitektur dasar IndoBERT untuk menangani tugas klasifikasi tiga kelas. Penggunaan data validasi selama proses pelatihan berfungsi untuk memantau *loss* dan akurasi, guna mencegah terjadinya *overfitting*.

Hasil Evaluasi Model

Kinerja model dalam mengklasifikasikan sentimen diukur menggunakan metrik standar seperti *accuracy, precision, recall,* dan *F1-score*. Evaluasi ini memberikan gambaran objektif mengenai sejauh mana model IndoBERT dapat diandalkan dalam memetakan persepsi publik. Untuk melihat detail prediksi model pada tiap label, dilakukan analisis menggunakan *Confusion Matrix*. Gambar 4 menunjukkan perbandingan antara label asli (*True Label*) dan hasil prediksi model (*Predicted Label*) pada data uji



Gambar 4. *Confusion Meatrix*

Berdasarkan Gambar 4, model menunjukkan akurasi yang tinggi dengan mayoritas data terakumulasi pada garis diagonal. Tercatat 454 data netral, 196 data negatif, dan 43 data positif berhasil diklasifikasikan dengan benar. Menariknya, tidak ditemukan kesalahan klasifikasi antara sentimen positif dan negatif (bernilai 0), yang membuktikan kemampuan model dalam membedakan polaritas sentimen secara tegas.

Selanjutnya, performa model diukur secara lebih rinci menggunakan *Classification Report*. Tabel pada Gambar 5 menyajikan nilai *precision, recall, dan f1-score* untuk setiap kategori sentimen:

Classification Report (Test Set):

	precision	recall	f1-score	support
negatif	0.9800	0.9561	0.9679	205
netral	0.9722	0.9891	0.9806	459
positif	0.9773	0.9149	0.9451	47
accuracy			0.9747	711
macro avg	0.9765	0.9534	0.9645	711
weighted avg	0.9748	0.9747	0.9746	711

Gambar 5. *Classification Report*

Gambar 5 menunjukkan nilai *f1-score* yang sangat stabil di angka 0,94 hingga 0,98 untuk seluruh kelas. Model mencapai akurasi keseluruhan sebesar 97,47%, dengan performa tertinggi pada kelas netral (*f1-score* 0,98). Nilai *weighted average* yang konsisten di angka 0,97 membuktikan bahwa meskipun terdapat perbedaan jumlah data (*support*), model tetap mampu memberikan prediksi yang akurat dan seimbang di semua kategori sentimen

Berdasarkan hasil evaluasi tersebut, model IndoBERT mencapai tingkat akurasi keseluruhan sebesar 97,47%. Angka yang sangat tinggi ini membuktikan bahwa mekanisme *self-attention* pada IndoBERT bekerja sangat efektif dalam memahami struktur kalimat masyarakat Indonesia di media sosial. Skor tinggi pada *precision* dan *recall* juga menunjukkan bahwa model memiliki kemampuan yang merata dalam mendeteksi sentimen positif, negatif, maupun netral tanpa terjadi ketimpangan performa yang signifikan.

PEMBAHASAN

Pembahasan Pra-processing Data

Hasil pra-pemrosesan menunjukkan bahwa elemen nonteks seperti emoji dan tanda baca berlebihan berhasil dihilangkan, huruf kapital diseragamkan, serta kata tidak baku dinormalisasi ke dalam bentuk Bahasa Indonesia yang lebih standar. Proses ini menghasilkan teks yang lebih bersih dan konsisten, sehingga memudahkan model dalam memahami konteks kalimat pada tahap klasifikasi sentimen.

Pembahasan Distribusi Sentimen

Distribusi label sentimen menunjukkan adanya ketidakseimbangan data, dengan dominasi kelas netral. Hal ini mengindikasikan bahwa sebagian besar komentar bersifat informasional dan tidak mengekspresikan sikap emosional yang ekstrem. Kondisi ini mencerminkan kecenderungan publik untuk menyampaikan pertanyaan, klarifikasi, atau pernyataan faktual terkait isu yang dibahas.

Analisis Word Cloud Sentimen Negatif

Gambar 4. *Word cloud* Negatif

Visualisasi word cloud pada sentimen negatif menunjukkan dominasi kata-kata seperti *"tidak"*, *"rakyat"*, *"DPR"*, dan *"perampasan"*. Pola ini mengindikasikan adanya ekspresi ketidakpuasan publik terhadap kebijakan maupun sikap DPR.Kritik publik terbagi ke dalam dua kecenderungan utama, yaitu kritik substantif terhadap ketidakjelasan regulasi dan kritik normatif terhadap kinerja serta responsivitas wakil rakyat.

Analisis Word Cloud Sentimen Netral

Gambar 5. *Word cloud* Netral

Word cloud sentimen netral memperlihatkan dominasi kata-kata seperti *"rakyat"*, *"DPR"*, *"benar"*, dan *"kalau"*. Hal ini menunjukkan bahwa komentar netral cenderung berorientasi pada pencarian informasi dan klarifikasi. Sikap ini mencerminkan kehati-hatian publik dalam merespons isu sensitif sebelum membentuk penilaian yang lebih tegas.

Pembahasan Kinerja Model

Berdasarkan confusion matrix dan classification report, model IndoBERT menunjukkan performa yang sangat baik dengan akurasi 97,47%. Model mampu mengklasifikasikan sentimen netral dengan ketepatan tertinggi, diikuti oleh sentimen negatif dan positif. Nilai precision, recall, dan F1-score yang relatif seimbang menunjukkan bahwa model tidak mengalami bias signifikan terhadap kelas tertentu, meskipun data tidak

seimbang. Hal ini menegaskan stabilitas dan keandalan IndoBERT dalam klasifikasi sentimen berbahasa Indonesia.

KESIMPULAN

Berdasarkan hasil analisis yang telah dilakukan, penerapan model IndoBERT terbukti sangat efektif dalam memetakan persepsi masyarakat terhadap dinamika sosial-politik nasional di platform digital dengan capaian akurasi sebesar **97,47%** serta nilai metrik evaluasi yang seimbang. Keberhasilan ini menunjukkan bahwa arsitektur berbasis *Transformer* mampu menangkap kompleksitas bahasa informal, penggunaan *slang*, dan nuansa semantik pada komentar Instagram secara presisi meskipun data bersifat tidak terstruktur. Temuan penelitian mengungkapkan bahwa opini publik didominasi oleh sentimen netral yang mencerminkan kecenderungan masyarakat untuk bersikap hati-hati dan informasional, sementara sentimen negatif muncul sebagai representasi kritik terhadap kebijakan lembaga terkait, dan sentimen positif merujuk pada apresiasi atas transparansi informasi. Dengan demikian, penggunaan IndoBERT dalam kerangka kerja *Knowledge Discovery in Databases* (KDD) dapat diandalkan sebagai instrumen strategis untuk memantau dinamika opini publik secara sistematis dan berbasis data, guna mendukung proses evaluasi kebijakan yang lebih responsif di Indonesia.

REFERENSI

Asri, Y., Kuswardani, D., Ramadhana, S. A., TS, J. F. P., Marbun, D. U. N., Fatimah, F. N., & Qoriza, Z. (2025). *OPTIMALISASI ANALISIS SENTIMEN DENGAN SPELLING CORRECTOR*. Uwais Inspirasi Indonesia. https://books.google.co.id/books?id=WItmEQAAQBAJ

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). 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* (Vol. 1). https://doi.org/10.48550

Geni, L., Yulianti, E., & Sensuse, D. I. (2023). Sentiment Analysis of Tweets Before the 2024 Elections in Indonesia Using Bert Language Models. *Jurnal Ilmiah Teknik Elektro Komputer Dan Informatika*, *9*(3), 746–757. https://doi.org/10.26555/jiteki.v9i3.26490

Geron, A. (2022). *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow* (3rd ed.). O'Reilly Media.

Halberstadt, J., de Bronstein, A. A., Greyling, J., & Bissett, S. (2022). Transforming Entrepreneurship Education: Interdisciplinary Insights on Innovative Methods and Formats. In *Transforming Entrepreneurship Education: Interdisciplinary Insights on Innovative Methods and Formats*. https://doi.org/10.1007/978-3-031-11578-3

Khairani, U., Mutiawani, V., & Ahmadian, H. (2024). Pengaruh Tahapan Preprocessing Terhadap Model Indobert Dan Indobertweet Untuk Mendeteksi Emosi Pada Komentar Akun Berita Instagram. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, *11*(4), 887–894. https://doi.org/10.25126/jtiik.1148315

Koto, F. (2021). *I NDO BERT WEET : A Pretrained Language Model for Indonesian Twitter with Effective Domain-Specific Vocabulary Initialization*. 10660–10668.

Kumar, A., Rani, S., Kumar, K. D., & Jain, M. (2025). *Handbook of AI in Engineering Applications: Tools, Techniques, and Algorithms*. CRC Press. https://books.google.co.id/books?id=k8VoEQAAQBAJ

Prajarini, D. (2021). *Media Sosial Periklanan - Instagram*. Deepublish. https://books.google.co.id/books?id=txdPEQAAQBAJ

Sari, R. M., Rizka, A., Putri, N. A., & Efriana, A. (2024). *Perhitungan Metode Clustering*. Serasi Media Teknologi. https://books.google.co.id/books?id=RIU0EQAAQBAJ

Setiawan, A. (2025). "OPINI PUBLIK: Teori dan Praktek." In A. Junaid (Ed.), *Perkumpulan Rumah Cemerlang Indonesia All Right Reserved*.

Sheikh, J. A., Khan, T., & Kanaujia, B. K. (2024). *Intelligent Signal Processing and RF Energy Harvesting for State of art 5G and B5G Networks*. Springer Nature Singapore. https://books.google.co.id/books?id=l2P8EAAAQBAJ

Suparwito, H., Gunawan, R., Binanto, I., Widyastuti, W., Kumalasanti, R. A., M, T. A. H., Nugroho, S. C., & Press, S. D. U. (2024). *Pengantar Manajemen Data*. Sanata Dharma University Press. https://books.google.co.id/books?id=0aQ4EQAAQBAJ

Wikipedia. (2025). *17+8 Tuntutan Rakyat*. https://id.wikipedia.org/wiki/17%2B8_Tuntutan_Rakyat#Perumusan

Yessy Asri, S. T. M., & Dr. Dra. Dwina Kuswardani, M. K. (2024). *Text Similarity*. Uwais Inspirasi indonesia . https://books.google.co.id/books?id=YO7uEAAAQBAJ