JURNAL LOCUS: Penelitian & Pengabdian
Volume 5, No. 3, Maret 2026
E-ISSN 2829-7334| P-ISSN 2829-5439
Hompage: https://locus.rivierapublishing.id/index.php/jl

Analisis Sentimen Warganet terhadap Program Makan Bergizi Gratis (MBG) Menggunakan Model Indobert

M. Haikal[1]*, Adhitia Erfina2
Universitas Nusa Putra, Indonesia
Email: m.haikal_si22@nusaputra.ac.id*, adhitia.erfina@nusaputra.ac.id

Abstrak

Penelitian ini bertujuan untuk mengidentifikasi sentimen warganet terhadap Program Makan Bergizi Gratis (MBG) melalui analisis teks pada platform Twitter/X. Sebanyak 361 tweet dikumpulkan melalui proses scraping dan dipreproses melalui beberapa tahap, termasuk case folding, pembersihan teks, normalisasi, tokenisasi BERT, serta padding dan truncation. Model IndoBERT digunakan sebagai model utama untuk klasifikasi sentimen dan dilatih kembali (fine-tuning) menggunakan dataset berlabel tiga kategori, yaitu positif, negatif, dan netral. Hasil eksperimen menunjukkan bahwa model IndoBERT mampu melakukan klasifikasi dengan tingkat akurasi sebesar 60,27%, dengan nilai precision, recall, dan F1-score yang relatif seimbang. Distribusi sentimen memperlihatkan bahwa opini negatif sedikit lebih dominan dibandingkan sentimen positif dan netral. Sementara itu, confusion matrix mengindikasikan bahwa kelas netral merupakan kategori yang paling sulit dibedakan oleh model. Secara keseluruhan, penelitian ini menunjukkan bahwa IndoBERT dapat menjadi pendekatan yang efektif dalam menganalisis sentimen kebijakan publik di media sosial, meskipun kinerja model masih dapat ditingkatkan melalui perluasan dataset dan pelabelan manual yang lebih akurat.

Kata Kunci: Analisis Sentimen, BERT, MBG, Twitter/X,

Corresponding: M. Haikal
E-mail: m.haikal_si22@nusaputra.ac.id

PENDAHULUAN

Perkembangan teknologi informasi telah mempermudah masyarakat dalam menyampaikan pendapat melalui media sosial (Boyd & Crawford, 2019). Salah satu platform yang sering digunakan untuk mengekspresikan opini publik adalah Twitter/X, karena memungkinkan pengguna untuk berinteraksi dan memberikan tanggapan terhadap isu-isu sosial secara cepat dan terbuka (Pak & Paroubek, 2010). Data dari We Are Social (2024) mencatat bahwa Indonesia memiliki 24 juta pengguna aktif Twitter/X, menjadikannya salah satu negara dengan pengguna terbesar di dunia. Hal ini menjadikan Twitter/X sebagai sumber data yang kaya untuk menganalisis opini publik terhadap berbagai kebijakan pemerintah.

Salah satu topik yang sering diperbincangkan di Indonesia adalah Program Makan Bergizi Gratis (MBG) yang digagas oleh pemerintah sebagai bagian dari upaya meningkatkan gizi anak sekolah. Program ini menimbulkan beragam reaksi dari masyarakat, baik yang mendukung maupun yang mengkritik kebijakan program tersebut. Oleh karena itu, penting untuk mengetahui bagaimana sentimen warganet terhadap program MBG agar pemerintah dan pihak terkait dapat memahami persepsi publik secara objektif (Bakshi et al., 2020).

Analisis sentimen merupakan teknik yang digunakan untuk mengidentifikasi dan mengklasifikasikan opini publik menjadi positif, negatif, atau netral berdasarkan data teks (Liu, 2020). Penelitian sebelumnya tentang analisis sentimen kebijakan publik di Indonesia telah dilakukan dengan berbagai pendekatan. Rahman dan Wibowo (2023) menganalisis sentimen kebijakan subsidi BBM menggunakan Support Vector Machine (SVM) dan mencapai akurasi 75% (Rahman & Wibowo, 2023). Sari dan Gunawan (2023) menggunakan metode lexicon-based untuk menganalisis sentimen kebijakan pendidikan. Yuliani dan Prasetyo (2024) serta Yunitasari dan Hasanah (2022) telah menerapkan model BERT untuk analisis sentimen program Kartu Prakerja dan kebijakan PPKM. Namun, penelitian-penelitian tersebut belum ada yang secara spesifik menganalisis sentimen terhadap Program Makan Bergizi Gratis (MBG). Selain itu, penggunaan model BERT khusus bahasa Indonesia seperti IndoBERT untuk analisis sentimen program pemerintah masih relatif terbatas. Celah penelitian (research gap) inilah yang menjadi fokus utama penelitian ini.

Dalam penelitian ini, digunakan model BERT (Bidirectional Encoder Representations from Transformers) yang merupakan salah satu model deep learning terkini dalam bidang Natural Language Processing (NLP) (Jurafsky & Martin, 2023; Kumar & Jaiswal, 2023; Mahendra et al., 2021; Minaee et al., 2021; Vaswani et al., 2017). Berbeda dengan model sebelumnya seperti SVM atau LSTM, BERT memiliki kemampuan memahami konteks kata secara dua arah (bidirectional) sehingga hasil analisis sentimen lebih akurat (Koto et al., 2020; Medhat et al., 2014 Putri et al., 2022). Arsitektur BERT yang diperkenalkan oleh Devlin et al. (2019) didasarkan pada mekanisme attention yang memungkinkan model menangkap hubungan antar kata dalam kalimat secara lebih mendalam. Untuk konteks bahasa Indonesia, digunakan IndoBERT yang dikembangkan oleh Wilie et al. (2020) melalui proses pre-training pada korpus besar berbahasa Indonesia, sehingga lebih peka terhadap nuansa bahasa dan konteks lokal.

Bagian ini bertujuan untuk mengetahui persepsi warganet terhadap Program Makan Bergizi Gratis (MBG) yang ditunjukkan melalui berbagai unggahan di platform X. Selain itu, penelitian ini juga bertujuan mengimplementasikan model BERT sebagai alat analisis sentimen otomatis untuk mengklasifikasi teks secara lebih akurat. Hasil analisis tersebut kemudian disajikan dalam bentuk proporsi sentimen positif, negatif, dan netral sehingga dapat memberikan gambaran mengenai kecenderungan opini publik terhadap Program MBG.

Manfaat penting dari sisi akademik, penelitian ini dapat menjadi referensi dalam pengembangan kajian Natural Language Processing (NLP), khususnya terkait penerapan model BERT untuk bahasa Indonesia. Dari sisi praktis, hasil penelitian ini dapat dimanfaatkan oleh pemerintah untuk memahami bagaimana masyarakat menilai Program Makan Bergizi Gratis (MBG), sehingga kebijakan yang dirancang ke depan dapat lebih tepat sasaran. Selain itu, dari sisi teknologis, penelitian ini menunjukkan bagaimana teknologi kecerdasan buatan dapat digunakan untuk membaca dan memahami opini publik secara cepat dan efisien melalui data media sosial.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode analisis sentimen berbasis machine learning. BERT dipilih karena kemampuannya memahami konteks kalimat secara dua arah sehingga lebih akurat dalam memprediksi sentimen dibanding pendekatan berbasis kamus atau algoritma klasik.

Pengumpulan Data

Data dikumpulkan dari platform X menggunakan website Scraper dengan kata kunci yang relevan dengan Program Makan Bergizi Gratis (MBG), berbahasa Indonesia, dan jenis tweet terbaru.

Pembersihan dan Pemrosesan Teks

Tahapan ini dilakukan agar data siap diproses oleh model BERT. Langkah-langkahny meliputi:

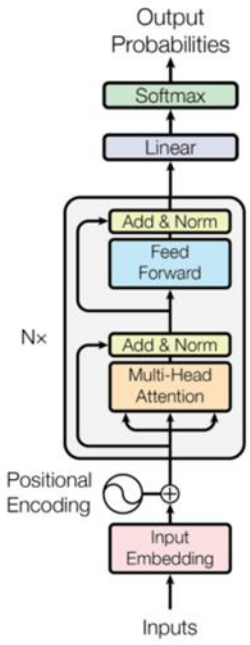
1. Case Folding, untuk mengubah teks menjadi huruf kecil.
2. Menghapus URL, mention, hastag, dan simbol yang tidak relevan.
3. Normalisasi kata gaul atau slang.
4. Tokenisasi BERT menggunakan WordPiece tokenizer.
5. Padding dan truncation, bertujuan agar panjang input seragam.

Pelabelan Sentimen

Setiap tweet yang telah dibersihkan diberi label sentimen positif, negatif, atau netral. Proses pelabelan dapat dilakukan secara manual oleh peneliti untuk menjaga akurasi, atau dengan bantuan model otomatis yang kemudian dilakukan verifikasi ulang. Label inilah yang nantinya digunakan sebagau dasar pelatihan model.

Model BERT dan Aritektur

Model yang digunakan dalam penelitian ini adalah BERT ynag telah melalui proses pre-training pada bahasa Indonesia, seperti IndoBERT. Model ini kemudian dilakukan fine-tuning agar lebih peka terhadap pola dan bahasa yang muncul pada data terkait MBG. BERT bekerja dengan memahami konteks dari dua arah, sehingga dapat menangkap makna yang lebih mendalam dari sebuah kalimat.



Gambar 1. Arsitektur BERT

Gambar 2 menunjukan arsitektur dasar BERT yang dibangun sepenuhnya di atas encoder dari model Transformer. Setiap lapisan encoder terdiri dari dua komponen utama, yaitu Multi-Head Self-Attention yang memungkinkan model memahami hubungan antar kata dalam satu kalimat secara konteks menyeluruh, serta Feed Forward Network yang memproses representasi hasil perhatian secara lebih dalam. Setiap blok dilengkapi dengan mekanisme Add & Norm untuk menjaga stabilitas pembelajaran. Input teks terlebih dahuku melalui tahap Input Embedding dan Positional Encoding sebelum diproses oleh tumpukan encoder berulang (Nx). Hasil akhir dari lapisan encoder kemudian diteruskan ke lapisan linear dan fungsi Softmax untuk menghasilkan klasifikasi, misalnya kategori sentimen positif, negatif, atau netral.

Pelatihan Model (Fine-tuning)

Dataset yang sudah diberi label dibagi menjadi data latih dan data uji. Proses fine-tuning dilakukan dengan menyesuaikan beberapa parameter seperti jumlah epoch, learning rate, dan ukuran batch (Sun et al., 2019). Model dilatih untuk mengenali pola sentimen berdasarkan data latih, kemudian diuji performanya menggunakan data uji.

Evaluasi Model

Evaluasi dilakukan untuk mengetahui sejauh mana model mampu melakukan klasifikasi sentimen secara akurat. Metrik yang digunakan dalam evaluasi meliputi:

1. Accuracy
2. Precision
3. Recall
4. F1-score

Hasil pengujian kemudian dibandingkan dan dianalisis untuk mengetahui performa model.

Tabel 1. Confusion Matrix

Class	Classified as Positive	Classified as Negative
Positive	True Positive (TP)	False Negative (FN)
Negative	False Positive (FP)	True Negative (TN)

Analisis Hasil

Tahap terakhir adalah menganalisis hasil evaluasi model dan pola sentimen yang muncul pada dataset. Analisis ini digunakan untuk melihat kecenderungan opini publik terhadap program Makan Bergizi Gratis, pakah didominasi sentimen positif, negatif, atau netral. Temuan pada tahap ini juga mrnjasdi dasar penyusunan kesimpulan dan rekomendasi pada bab berikutnya.

HASIL DAN PEMBAHASAN

Pengumpulan Sumber Data

Data penelitian berasal dari platform X dengan kata kunci “makan bergizi gratis”. Proses pengambilan data dilakukan menggunakan website Scraper. Dari proses awal, diperoleh sejumlah tweet mentah dengan total data yang diperoleh adalah 1000

Preprocessing

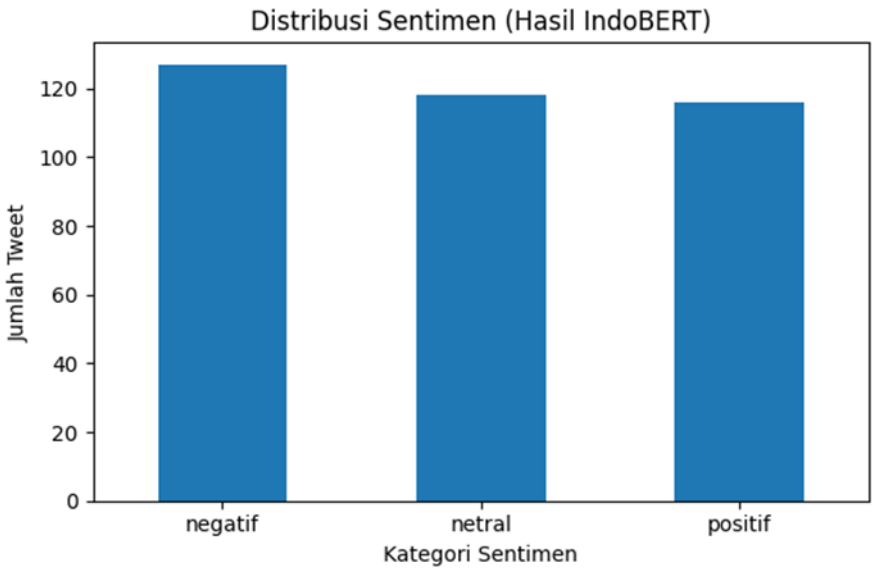
Kemudian dilakukan tahapan preprocessing yaitu proses pembersihan, case folding, Menghapus URL, mention, hastag, dan simbol yang tidak relevan, Normalisasi kata gaul atau slang, Tokenisasi, Padding dan truncation. Selain itu dilakukan penghapusan data duplikat hasil dari scraping.

Tabel 2. Contoh data setelah preprocessing

Preprocessing Text	Contoh Data
Original Data	Maaf sebelumnya, jika melihat keadaan saat ini, program MBG lebih baik dialihkan menjadi program kerja mudah atau ciptakan lapangan kerja yg lebih banyak. Jika lapangan kerja banyak maka InsyaAllah kebutuhan keluarga akan terpenuhi tercukupi.. #usul",https://x.com/andaru1108/status/1896705629932380559
Case Folding	maaf sebelumnya, jika melihat keadaan saat ini, program mbg lebih baik dialihkan menjadi program kerja mudah atau ciptakan lapangan kerja yg lebih banyak. jika lapangan kerja banyak maka insyaallah kebutuhan keluarga akan terpenuhi tercukupi.. #usul",https://x.com/andaru1108/status/1896705629932380559
Menghapus URL, mention, hastag, dan simbol yang tidak relevan	maaf sebelumnya jika melihat keadaan saat ini program mbg lebih baik dialihkan menjadi program kerja mudah atau ciptakan lapangan kerja yg lebih banyak. jika lapangan kerja banyak maka insyaallah kebutuhan keluarga akan terpenuhi tercukupi
Normalisasi kata	maaf sebelumnya jika melihat keadaan saat ini program mbg lebih baik

gaul atau slang,	dialihkan menjadi program kerja mudah atau ciptakan lapangan kerja yg lebih banyak. jika lapangan kerja banyak maka insyaallah kebutuhan keluarga akan terpenuhi tercukupi
Tokenisasi	['maaf', 'sebelumnya', 'jika', 'melihat', 'keadaan', 'saat', 'ini', 'program', 'mb', '##g', 'lebih', 'baik', 'dialihkan', 'menjadi', 'program', 'kerja', 'mudah', 'atau', 'ciptakan', 'lapangan', 'kerja', 'yg', 'lebih', 'banyak', '.', 'jika', 'lapangan', 'kerja', 'banyak', 'maka', 'insyaallah', 'kebutuhan', 'keluarga', 'akan', 'terpenuhi', 'terc', '##ukupi']
Padding dan *truncation*	Token ID setelah padding: tensor([2, 2727, 1131, 338, 722, 1762, 305, 92, 986, 2201, 30365, 216, 342, 16261, 234, 986, 494, 783, 158, 14018]) Attention Mask: tensor([1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1]) Total panjang token setelah padding/truncation: 64

Pelabelan Sentimen



Gambar 2. Visualisasi distribusi label sentimen

Pelabelan dilakukan menggunakan model IndoBERT yang telah melalui proses fine-tuning pada dataset penelitian ini. Pelabelan sentimen dilakukan untuk menentukan kategori opini yang terkandung dalam setiap tweet. Terdapat tiga kelas sentimen yang digunakan, yaitu negatif, netral, dan positif. Setiap sentimen kemudian dipetakan kedalam label numerik untuk kebutuhan pemodelan:

1. Negatif = 0
2. Netral = 1
3. Positif = 2

Berdasarkan hasil pelabelan otomatis terhadap seluruh data yang berjumlah 361 tweet, diperoleh distribusi sentimen dengan rincian 118 kategori netral, 116 kategori positif, dan 127

kategori negatif. Jumlah ini kemudian divisualisasikan dalm bentuk diagram batang untuk memberikan gambaran proporsi sentimen warganet terhadap program MBG sebagaiman yang ditampilkan pada gambar 2.

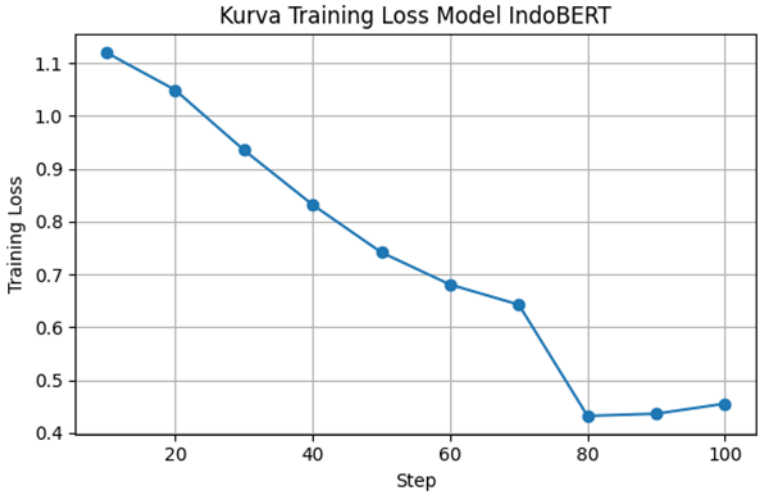
Pelatihan Model IndoBERT

Setelah proses pelabelan selesai, model IndoBERT dilatih menggunakan dataset yang sudah berlabel. Fine-tuning dilakukan menggunakan parameter seperti pada Tabel 4.2

Tabel 3. Hyperparameter fine-tuning IndoBERT

Hyperparameter	Nilai
Epoch	3
Batch size	8
Learning rate	2e-5
Optimizer	AdamW
Rasio data latih validasi	80% : 20%

Selama proses pelatihan, nilai training loss terus menurun seiring bertambahnya jumlah iterasi. Hal ini menunjukkan bahwa model berhasil mempelajari pola data dengan baik. Dari hasil training dicatat bahwa pada akhir epoch ketiga, nilai training loss berada pada kisaran $0.71 \rightarrow 0.45$, yang menunjukkan adanya konvergensi.



Gambar 3. Kurva training loss model IndoBERT

Evaluasi Model

Tabel 4. Hasil evaluasi model

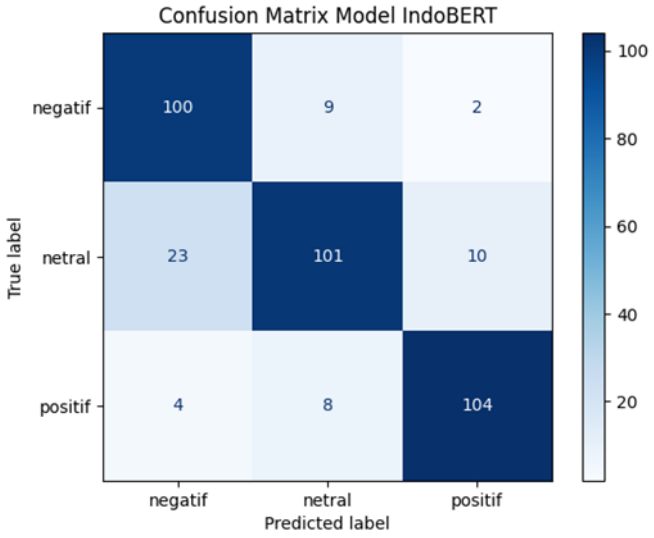
Metrik	Nilai
Accuracy	0.6027
Precision	0.6023
Recall	0.6027
F1-Score	0.5976

Hasil evaluasi model IndoBERT yang ditunjukkan pada Tabel 4.X mencerminkan bahwa model memiliki nilai akurasi sebesar 60,27% dengan nilai precision, recall, dan F1-score yang relatif seimbang pada kisaran 0,59–0,60. Keseragaman nilai pada keempat metrik tersebut menunjukkan bahwa model tidak bias terhadap salah satu kategori sentimen dan mampu melakukan generalisasi secara moderat terhadap data validasi.

Nilai akurasi 60% ini dapat dikategorikan cukup baik mengingat beberapa karakteristik dataset. Pertama, jumlah data relatif kecil, yaitu hanya 360 tweet, sehingga informasi yang dapat dipelajari model menjadi terbatas. Kedua, tweet bersifat pendek dan informal, sehingga konteks sentimen terkadang ambigu, khususnya pada sentimen netral yang sering kali berisi pernyataan tanpa ekspresi emosional yang jelas. Ketiga, pelabelan menggunakan pseudo-label berpotensi menghasilkan noise, sehingga memengaruhi kualitas sinyal pembelajaran saat fine-tuning.

Selain itu, hasil confusion matrix menunjukkan bahwa kesalahan klasifikasi terbesar terjadi pada kelas netral, terutama ketika tweet yang sebenarnya netral diprediksi sebagai negatif. Hal ini menunjukkan bahwa perbedaan semantik antara komentar netral dan komentar bernada kritik ringan cenderung sulit dipisahkan oleh model. Sebaliknya, kinerja model terhadap kelas positif dan negatif cenderung lebih baik. Hal ini dapat dilihat dari jumlah prediksi benar yang cukup tinggi pada kedua kelas tersebut.

Berdasarkan hasil evaluasi ini, dapat disimpulkan bahwa model IndoBERT sudah mampu menangkap pola sentimen dasar pada pembahasan publik mengenai program MBG, meskipun performanya masih dapat ditingkatkan dengan memperbesar jumlah data pelatihan, memperbaiki kualitas label, dan menambah jumlah epoch pelatihan.



Confusion matrix digunakan untuk melihat kemampuan model dalam membedakan setiap kelas sentiment. Akurasi tertinggi terdapat pada kelas negatif dan positif, dengan jumlah prediksi benar masing-masing 100 dan 104. Kelas netral paling sulit diprediksi,

ditandai tingginya misclassfication ke kelas negatif (23 data). Model menunjukkan kemampuan seimbang antara ketiga kelas, meskipun perbedaan semantik antara netral vs positif/negatif memerlukan lebih banyak data untuk meningkatkan akurasi. Misclass terbesar terdapat pada kelas Netral → Negatif (23 data), menunjukkan bahwa model masih sulit membedakan opini netral dari opini bernada kritik ringan

KESIMPULAN

Berdasarkan penelitian yang telah dilakukan mengenai analisis sentimen warganet terhadap Program Makan Bergizi Gratis (MBG) menggunakan model IndoBERT, dapat disimpulkan bahwa opini publik mengenai program tersebut bersifat beragam dengan kecenderungan sentimen negatif yang sedikit lebih dominan dibandingkan sentimen netral dan positif. Dari total 361 tweet yang dianalisis, diperoleh hasil bahwa 127 tweet termasuk dalam kategori negatif, 118 tweet netral, dan 116 tweet positif. Distribusi ini menunjukkan bahwa masyarakat memiliki pandangan yang cukup kritis terhadap MBG, namun masih terdapat proporsi yang signifikan dari pengguna yang memberikan opini netral maupun positif.

Model IndoBERT yang digunakan dalam penelitian mampu melakukan klasifikasi sentimen dengan performa yang cukup moderat. Setelah melalui proses fine-tuning selama tiga epoch, model menghasilkan nilai accuracy sebesar 60,27%, precision sebesar 60,23%, recall sebesar 60,27%, serta F1-score sebesar 59,76%. Performa ini menunjukkan bahwa IndoBERT dapat mengenali pola sentimen dengan cukup baik meskipun dataset yang dilatih relatif terbatas. Kurva training loss yang terus menurun selama proses pelatihan menegaskan bahwa model mampu menyesuaikan bobotnya secara stabil tanpa menunjukkan gejala overfitting.

Dari hasil confusion matrix, terlihat bahwa prediksi model paling akurat pada kelas negatif dan positif. Namun, kelas netral merupakan kategori yang paling sulit dibedakan, di mana cukup banyak tweet netral diprediksi sebagai negatif maupun positif. Hal ini wajar mengingat karakteristik tweet netral yang cenderung ambigu dan tidak memiliki ekspresi emosional yang jelas. Secara umum, hasil penelitian ini menunjukkan bahwa IndoBERT dapat digunakan sebagai alat bantu yang cukup efektif dalam menganalisis persepsi publik terhadap suatu isu kebijakan, termasuk program MBG.

Meskipun demikian, penelitian ini memiliki beberapa keterbatasan, terutama pada ukuran dataset yang kecil dan penggunaan pseudo-label pada tahap awal pelabelan. Oleh karena itu, penelitian selanjutnya disarankan untuk menggunakan dataset yang lebih besar serta melakukan pelabelan manual untuk meningkatkan kualitas data. Selain itu, penggunaan model bahasa lain seperti IndoBERTweet atau RoBERTa-Indonesia juga dapat dieksplorasi untuk memperoleh performa yang lebih baik. Eksperimen lanjutan dengan variasi hyperparameter dan jumlah epoch yang lebih banyak juga berpotensi memberikan hasil yang lebih optimal. Secara keseluruhan, penelitian ini memberikan gambaran yang komprehensif mengenai persepsi warganet terhadap program MBG dan menunjukkan bahwa model IndoBERT dapat menjadi alternatif yang andal dalam analisis sentimen berbahasa Indonesia.

REFERENSI

Bakshi, R. K., Kaur, N., & Kaur, R. (2020). Opinion mining and sentiment analysis using neural networks. Procedia Computer Science, 173, 312–321.

Boyd, D., & Crawford, K. (2019). Critical questions for big data: Challenges for social media research. Information, Communication & Society, 15(5), 662–679.

Cahyani, R. N., & Adriani, M. (2021). Neural machine translation for Indonesian-English using Transformer architecture. Journal of Big Data, 8(1), 1–15.

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In Proceedings of NAACL-HLT (pp. 4171–4186).

Jurafsky, D., & Martin, J. H. (2023). Speech and language processing (3rd ed.). Pearson.

Koto, F., Lau, J. H., & Baldwin, T. (2020). IndoBERTweet: A pretrained language model for Indonesian social media text. In Proceedings of the 2020 Workshop on Asian Language Resources (pp. 17–25).

Kumar, A., & Jaiswal, A. (2023). Text preprocessing techniques in natural language processing. International Journal of Artificial Intelligence Research, 7(2), 55–64.

Liu, B. (2020). Sentiment analysis: Mining opinions, sentiments, and emotions. Cambridge University Press.

Mahendra, R., Rahmad, K., Awang, A., & Sari, N. (2021). IndoLEM: A benchmark dataset for Indonesian NLP tasks. In Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics.

Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. Ain Shams Engineering Journal, 5(4), 1093–1111.

Minaee, S., Kalchbrenner, N., Cambria, E., Nikzad, N., Chenaghlu, M., & Gao, J. (2021). Deep learning–based text classification: A comprehensive review. ACM Computing Surveys, 54(3), 1–40.

Pak, A., & Paroubek, P. (2010). Twitter as a corpus for sentiment analysis and opinion mining. In Proceedings of the International Conference on Language Resources and Evaluation (LREC) (pp. 1320–1326).

Pratama, Y., & Nugraha, R. (2021). Sentiment analysis on public opinion of government policies using BERT and SVM. Indonesian Journal of Artificial Intelligence and Data Mining, 4(2), 102–109.

Putri, D. A., Kurniawan, S., & Pratama, Y. (2022). Sentimen pengguna terhadap layanan e-commerce menggunakan Long Short-Term Memory. Jurnal Teknologi Informasi dan Komputer, 7(1), 45–53.

Rahman, A., & Wibowo, H. (2023). Analisis sentimen kebijakan subsidi BBM menggunakan Support Vector Machine. Jurnal Informatika Indonesia, 9(2), 112–120.

Sari, M., & Gunawan, A. (2023). Analisis sentimen kebijakan pendidikan menggunakan metode lexicon-based. Jurnal Data Mining Indonesia, 5(3), 88–95.

Sun, C., Qiu, X., Xu, Y., & Huang, X. (2019). How to fine-tune BERT for text classification? In China National Conference on Chinese Computational Linguistics (pp. 194–206).

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... &

Polosukhin, I. (2017). Attention is all you need. In Advances in Neural Information Processing Systems (pp. 5998–6008).

Wilie, B., Vincentio, K., & Putra, P. (2020). IndoBERT: A pretrained language model for Indonesian NLP tasks. In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP).

Yuliani, S., & Prasetyo, D. (2024). Analisis sentimen program Kartu Prakerja menggunakan IndoBERT. Jurnal Sistem Cerdas, 12(1), 33–41).

Yunitasari, F., & Hasanah, N. (2022). Analisis sentimen masyarakat terhadap kebijakan PPKM menggunakan BERT. Jurnal Teknologi Informasi dan Komputer, 8(1), 55–62.