Jurnal Algoritma, Logika dan Komputasi *Vol.VIII (No.02) : 820 - 829. Th. 2025*
Versi Online: https://journal.ubm.ac.id/index.php/alu *p-ISSN: 2620-620X*
DOI: http://dx.doi.org/10.30813/j-alu.v2i2.8868 *e-ISSN: 2621-9840*

Analisis Sentimen Program Makan Siang Gratis Menggunakan Model IndoBERT

[Sentiment Analysis of the Free Lunch Program Using the IndoBERT Model]

Jose Andreas Lukmanto, joseandreas31@gmail.com[1)], Theresia Puspa Wijayanti, twijayanti@bundamulia.ac.id[2)]
[1)2)] Program Studi Informatika/Fakultas Teknologi dan Desain, Universitas Bunda Mulia

Diterima 18 Juli 2025 / Disetujui 13 Agustus 2025

ABSTRACT

This research aims to analyze public sentiment regarding the free school lunch program in Indonesia utilizing the IndoBERT-Large-P2 model. The study used a dataset of 5,699 tweets, which were subsequently cleaned to 5,294 data points, to evaluate the model's performance in classifying positive, negative, and neutral sentiments toward the policy. The methodology encompassed data crawling using tweet-harvest, data preprocessing—including text cleaning and word normalization—and dataset splitting into training, validation, and testing sets. The IndoBERT-Large-P2 model was systematically tested with various configurations, including adjustments to batch size, learning rate, epochs, and data splits, to ascertain the optimal setup for sentiment classification. The analysis indicates that the model achieved a peak accuracy of 80% and a mean AUC score of 91%. This optimal performance was recorded with a configuration of batch size 8, a learning rate of 3e−5, an epsilon of 1e−9, training for 7 epochs, and a data split ratio of 68% for training, 12% for validation, and 20% for testing. This study concludes that the IndoBERT-Large-P2 model demonstrates robust performance in analyzing sentiment on the free lunch program on social media X, offering valuable insights into the application of advanced NLP models for public opinion analysis in Indonesia.
***Keywords**: Sentiment Analysis, IndoBERT-Large-P2, Free Lunch Program, Sentiment Classification, Social Media X*

ABSTRAK

Penelitian ini bertujuan untuk menganalisis sentimen publik terhadap program makan siang gratis di Indonesia menggunakan model IndoBERT-Large-P2. Penelitian ini menggunakan data *tweet* hasil *crawling* sebanyak 5,699 yang kemudian dibersihkan menjadi 5,294 data, dengan tujuan untuk mengevaluasi kinerja model dalam mengklasifikasikan sentimen positif, negatif, dan netral terhadap kebijakan tersebut. Metodologi yang digunakan meliputi tahapan *crawling* data menggunakan tweet-harvest, *preprocessing* data yang mencakup pembersihan teks dan normalisasi kata, serta pembagian data menjadi data latih, data validasi, dan data uji. Model IndoBERT-Large-P2 diuji dengan berbagai konfigurasi, termasuk pengaturan batch size, learning rate, epoch, dan pembagian data untuk memperoleh hasil terbaik dalam klasifikasi sentimen. Hasil analisis menunjukkan bahwa model dapat mencapai akurasi 80% dan nilai rata-rata AUC sebesar 91% dengan performa terbaik tercatat pada pengujian dengan *batch size* 8, *learning rate* 3e-5, *epsilon* 1e-9, *epoch* ke-7, dan rasio pembagian data 68% data latih, 12% data validasi, serta 20% data uji. Penelitian ini menyimpulkan bahwa model IndoBERT-Large-P2 memiliki kinerja yang baik dalam menganalisis sentimen terhadap program makan siang gratis di media sosial X dan memberikan wawasan penting mengenai penerapan model NLP dalam analisis opini publik di Indonesia.
Kata kunci: Analisis sentimen, IndoBERT-Large-P2, Program Makan Siang Gratis, Klasifikasi Sentimen, Media Sosial X

PENDAHULUAN

Pada era serba digital seperti sekarang, media sosial telah menjadi platform utama masyarakat untuk menyampaikan pendapat dan perasaan mereka mengenai berbagai topik. Secara definitif, media sosial merupakan jaringan di masyarakat yang bersifat inovatif dengan memanfaatkan teknologi dan aplikasi berbasis web yang dapat menciptakan peluang untuk

*Korespondensi Penulis:
E-mail: joseandreas31@gmail.com

membangun hubungan sosial dengan individu lain yang memiliki minat serupa [1]. Media sosial X merupakan salah satu platform yang banyak digunakan oleh masyarakat Indonesia untuk menyampaikan kritik maupun dukungan karena platform ini berfokus pada konten berbasis teks sehingga sering kali menjadi pusat perbincangan dan menghasilkan topik yang viral atau *trending* [2]. Hal ini dapat dibuktikan oleh hasil laporan dari We are Social yang menjadikan Indonesia sebagai target periklanan karena Indonesia memiliki 25,163 juta pengguna X aktif per Februari 2025 yang menjadikan Indonesia sebagai negara pengguna X terbesar nomor urut ke-3 di dunia [3].

Salah satu topik yang sering menjadi perbincangan masyarakat Indonesia saat ini adalah program makan siang gratis yang bertujuan untuk meningkatkan gizi anak dan ibu hamil, serta mengatasi masalah stunting di tanah air. Program makan siang gratis menjadi topik hangat di kalangan masyarakat Indonesia karena menjadi salah satu janji politik utama Prabowo Subianto dan Gibran Rakabuming Raka pada masa pemilihan umum. Mengingat keragaman opini publik terhadap kebijakan ini, penerapan analisis sentimen menjadi sangat penting. Analisis sentimen adalah sebuah tahapan yang bertujuan untuk memahami opini publik atau pengguna terhadap suatu topik atau platform dengan memanfaatkan data yang umumnya diperoleh dari internet serta media sosial, sehingga memungkinkan pengolahan dan interpretasi sentimen dalam skala luas guna mengidentifikasi pola serta kecenderungan dalam persepsi masyarakat [4]. Melalui metode ini, pemahaman yang lebih mendalam mengenai pandangan publik terhadap program pemerintah dapat diperoleh secara sistematis.

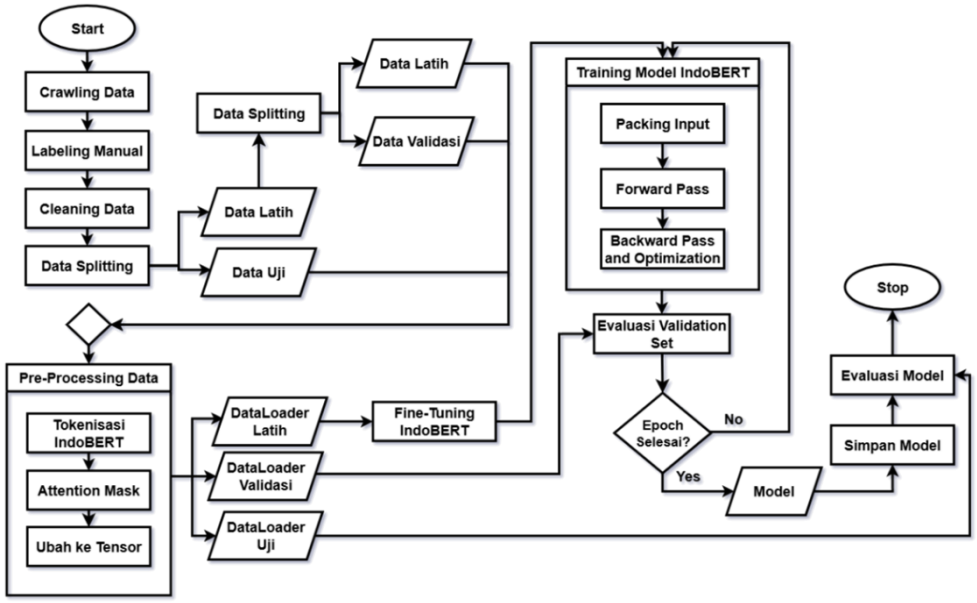
Pada penelitian sebelumnya yang dilakukan oleh Sitanggang, dkk [5], analisis sentimen terhadap program makan siang gratis dilakukan menggunakan model Naive Bayes. Penelitian tersebut menghasilkan akurasi sebesar 72,2%. Meskipun memberikan gambaran awal, model Naive Bayes memiliki keterbatasan fundamental karena bekerja berdasarkan frekuensi kata dan asumsi independensi antar kata [6]. Akibatnya, model ini kurang mampu menangkap makna berdasarkan konteks, urutan kata, maupun nuansa bahasa yang kompleks seperti ironi atau sarkasme yang sering ditemukan dalam opini di media sosial.

Namun seiring dengan perkembangan teknologi *Natural Language Processing* (NLP), muncullah berbagai model yang lebih canggih untuk analisis teks. NLP sendiri merupakan sebuah cabang AI yang memampukan komputer untuk memproses dan memahami bahasa manusia [7]. *Bidirectional Encoder Representations from Transformers* (BERT) merupakan salah satu dari model baru tersebut yang menarik untuk diteliti lebih lanjut. Model BERT ini dapat memahami konteks dari kata atau frasa dengan melakukan analisis secara *bidirectional* sehingga dapat memahami makna dari konteks secara lebih baik [8].

IndoBERT merupakan salah satu varian *pre-trained* BERT yang telah dilakukan pelatihan khusus untuk bahasa Indonesia dibandingkan model BERT standar yang berbasis bahasa Inggris, sehingga penggunaan model IndoBERT menjadi lebih baik dalam berbagai penggunaan tugas NLP yang menggunakan bahasa Indonesia. Hal ini dibuktikan dalam penelitian yang menggunakan model IndoBERT untuk melakukan analisis sentimen terhadap kinerja Penjabat Gubernur DKI Jakarta dengan data sebanyak 6.648 *tweet* yang dilakukan *preprocessing* dan pelabelan secara manual. Hasil dari model IndoBERT menunjukkan *accuracy* sebesar 90.5%, *precision* sebesar 90.6%, *recall* sebesar 90.5%, dan *f1-score* sebesar 90.49% pada data uji [9]. Penelitian lain juga menunjukkan bahwa IndoBERT memiliki kinerja yang unggul dalam analisis sentimen pada aplikasi layanan kesehatan, dengan data sebanyak 9310 dan dikategorikan menjadi 2 label yaitu positif dan negatif menghasilkan nilai akurasi mencapai 96%, *precision* sebesar 95%, *recall* sebesar 96%, dan *f1-score* sebesar 95% [10].

Oleh karena itu sasaran utama dari penelitian ini adalah untuk mengevaluasi secara sistematis kinerja dan efektivitas model IndoBERT-Large-P2 dalam tugas analisis sentimen tiga kelas (positif, negatif, netral) pada isu kebijakan publik yang dinamis seperti program makan siang gratis. Hubungan dengan penelitian terdahulu adalah penelitian ini secara langsung menguji apakah model yang lebih modern dan sadar konteks dapat memberikan peningkatan performa yang signifikan dibandingkan pendekatan probabilistik sederhana pada topik yang sama. Kebaharuan penelitian ini terletak pada penerapan dan optimisasi sistematis model *state-of-the-art* khusus Bahasa Indonesia pada analisis sentimen isu kebijakan publik yang relevan dan terkini, serta penyediaan analisis kinerja yang komprehensif.

METODE PENELITIAN



Gambar 1. Metode Penelitian

Penelitian ini terdiri dari beberapa tahapan utama yang dijelaskan berikut:

1. *Crawling Data*

Data dikumpulkan dari media sosial X menggunakan tweet-harvest dengan kata kunci "makan siang gratis". Rentang waktu ditetapkan dari 6 Januari 2025 hingga 5 Mei 2025, menghasilkan 5.699 data mentah. Periode ini dipilih untuk menangkap sentimen publik setelah program mulai direalisasikan, sehingga analisis berfokus pada reaksi terhadap implementasi nyata.

2. *Labeling* Manual

Tahapan selanjutnya setelah tahapan *crawling* data adalah *labeling* manual di mana peneliti akan memberikan label terhadap *tweet* yang telah dikumpulkan dengan kategori sentimen yang sesuai yaitu positif, negatif, maupun netral. *Labeling* manual ini sangat penting, terutama untuk meningkatkan akurasi model dalam mengidentifikasi sentimen, sekaligus membantu model dalam memahami aspek yang lebih kompleks seperti sarkasme yang sering kali tidak dapat ditangkap hanya dengan analisis berbasis kata kunci.

3. *Cleaning Data*
   Data yang telah dikumpulkan kemudian dilakukan beberapa tahap pembersihan:

   a. *Formatting Dataset* yaitu mengambil kolom full_text dan menambahkan kolom label, kemudian menyimpan hasilnya dalam CSV.
   b. *Text Cleaning* yaitu proses pembersihan untuk membersihkan dan mengubah data mentah menjadi format yang lebih terstruktur [11]. Pembersihan tersebut meliputi penghapusan URL, *mention* (@namauser), emoji, simbol, karakter khusus, dan angka, serta melakukan *case folding* untuk mengubah teks menjadi huruf kecil. Data yang telah dibersihkan disimpan dalam CSV.
   c. *Labeling* yaitu proses pelabelan dilakukan secara manual untuk 5,294 *tweet* untuk meningkatkan akurasi label.
   d. Normalisasi Teks menggunakan kamus kata baku untuk mengganti kata tidak baku dengan kata baku, kemudian menyimpan hasil normalisasi dalam CSV.

4. *Data Splitting*
   Data dibagi menjadi dua bagian yaitu 80% untuk *training* dan *validation set*, serta 20% untuk *test set*. Data *training* kemudian dibagi kembali dengan 85% untuk *training set* dan 15% untuk *validation set*.

5. *Pre-Processing Data*
   Data teks yang sudah bersih kemudian diproses agar sesuai dengan format *input* IndoBERT. Tahapan ini meliputi:

   a. Tokenisasi IndoBERT: Teks diubah menjadi token atau unit kata.
   b. Pembuatan Attention *Mask*: *Mask* dibuat untuk membedakan antara token asli dan *padding*.
   c. Konversi ke Tensor: Seluruh data seperti ID token, *attention mask*, dan label diubah menjadi format tensor PyTorch.

6. *Fine-tuning* IndoBERT
   Model IndoBERT-Large-P2 dilakukan *fine-tune* menggunakan algoritma optimasi AdamW dengan *learning rate* 3e-5 dan Epsilon 1e-9. Proses ini dilakukan dengan menggunakan GPU untuk efisiensi dan stabilitas.

7. *Training Model*
   Model dilatih dengan menggunakan *mini-batch gradient descent*. Proses ini mencakup *forward pass*, perhitungan *loss*, *backward pass*, dan optimasi menggunakan teknik seperti *gradient clipping* untuk mencegah *exploding gradient*.

8. Simpan Model
   Model yang telah dilatih disimpan dalam format direktori terstruktur menggunakan pustaka transformers, untuk digunakan kembali dalam inferensi atau pengujian lanjutan.
9. Evaluasi Model
   Evaluasi dilakukan dengan *test set* menggunakan metrik akurasi, ROC-AUC dan *confusion matrix* untuk mengukur kinerja model dalam klasifikasi sentimen.

METODE PENGUJIAN

Berbagai metode pengujian diperlukan untuk mengevaluasi kinerja model klasifikasi sentimen menggunakan model IndoBERT-Large-P2 pada data program makan siang gratis.

Evaluasi dilakukan dengan metrik ROC-AUC, *Confusion Matrix, Accuracy, Precision, Recall*, dan *f1-score*. Proses pengujian terdiri dari beberapa tahap:

1. Eksperimen *batch size* dan *epoch*

Pada tahap pertama, dilakukan eksperimen untuk menemukan kombinasi *batch size* dan jumlah *epoch* yang optimal. Pengujian ini menggunakan variasi *batch size* (8, 16, 32) dan dilatih selama 10 *epoch*. Konfigurasi *baseline* seperti learning rate 3e-5 dan epsilon 1e-9 dipilih karena merupakan nilai yang umum direkomendasikan dalam literatur untuk proses *fine-tuning* model berbasis BERT dan telah terbukti memberikan titik awal yang stabil pada berbagai tugas klasifikasi teks di penelitian-penelitian sebelumnya. Tujuan eksperimen ini adalah untuk menemukan kombinasi yang memberikan akurasi validasi tertinggi tanpa mengalami *overfitting*.

2. ROC AUC dan Confusion Matrix

Setelah mendapatkan konfigurasi terbaik dari tahap pertama, model dievaluasi secara menyeluruh pada *test set*. *Confusion Matrix* digunakan untuk menganalisis distribusi kesalahan klasifikasi secara detail [12]. yang kemudian diuraikan lebih lanjut dengan metrik *precision, recall,* dan f1-score melalui *classification report* [13].

Kurva ROC-AUC digunakan untuk mengukur kemampuan diskriminatif model, yaitu sejauh mana model dapat membedakan antara kelas sentimen yang satu dengan yang lainnya [14]. ROC curve menggambarkan hubungan antara *True Positive* Rate (TPR) atau sensitivitas dan *False Positive Rate* (FPR) yang menunjukkan *trade-off* antara kemampuan model dalam mengidentifikasi kasus positif dengan risiko salah mengklasifikasikan kasus negatif, sedangkan nilai AUC mewakili keseluruhan kemampuan diskriminatif model berdasarkan kurva ROC [15].

HASIL DAN PEMBAHASAN

Pada tabel 1, Tahap pengujian ini menggunakan rasio pembagian data 80:20 dengan *learning rate* 3e−5, dan *epsilon* 1e−9, pengujian dilanjutkan untuk mencari *batch size* dan jumlah *epoch* optimal.

Tabel 1. Implementasi *Batch Size* 8

Epoch	Loss	Akurasi
1	0.67	0.79
2	0.36	0.75
3	0.15	0.79
4	0.05	0.81
5	0.02	0.82
6	0.01	0.82
7	0.00	0.83
8	0.00	0.82
9	0.00	0.81
10	0.00	0.82

Pada Tabel 2, dengan *batch size* 8, model mencapai akurasi puncak tertinggi di antara semua pengujian, yaitu 0.83 pada *epoch* ke-7. Penurunan akurasi pada epoch ke-8 dan ke-9, merupakan indikator kuat bahwa model mulai *overfitting*. Oleh karena itu, epoch ke-7 adalah titik pemberhentian yang ideal.

Tabel 2. Implementasi *Batch Size* 16

Epoch	Loss	Akurasi
1	0.67	0.80
2	0.34	0.81
3	0.13	0.81
4	0.06	0.80
5	0.02	0.81
6	0.01	0.82
7	0.01	0.82
8	0.00	0.82
9	0.00	0.82
10	0.00	0.82

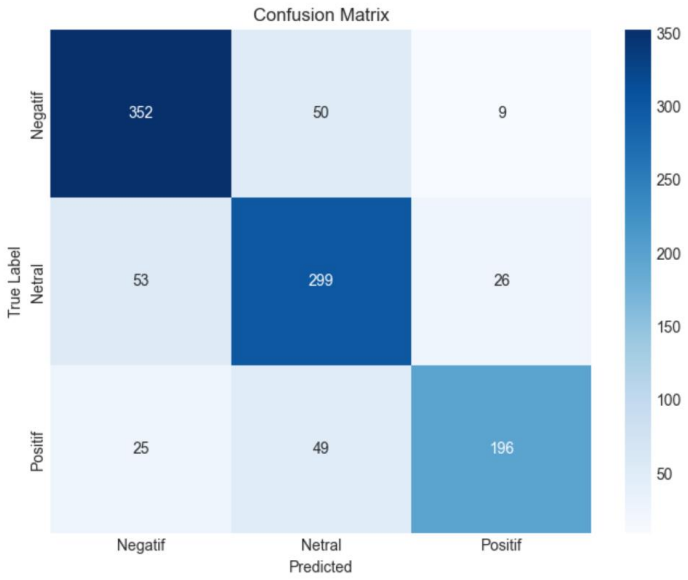
Pada Tabel 3, dengan *batch size* 16, model menghasilkan performa yang stabil dengan akurasi maksimal 0.82. Meskipun hasilnya konsisten, nilai akurasi puncaknya tidak melampaui yang dicapai oleh *batch size* 8. Stagnasi performa setelah *loss* mencapai nol juga terlihat di sini.

Tabel 3. Implementasi *Batch Size* 32

Epoch	Loss	Akurasi
1	0.71	0.80
2	0.39	0.80
3	0.17	0.80
4	0.06	0.81
5	0.03	0.80
6	0.02	0.81
7	0.00	0.81
8	0.00	0.81
9	0.00	0.81
10	0.00	0.81

Dengan *batch size* 32, performa model cenderung stagnan pada level akurasi 0.81. Fenomena *training loss* mencapai nol namun akurasi validasi tidak meningkat lebih lanjut menunjukkan bahwa *batch size* ini kurang efektif untuk menemukan performa optimal pada *dataset* ini.

Berdasarkan hasil tabel 3, konfigurasi dengan *batch size* 8 dan pelatihan selama 7 *epoch* dipilih sebagai yang terbaik. Kombinasi ini berhasil mencapai akurasi validasi tertinggi tepat sebelum model menunjukkan gejala *overfitting* yang signifikan, sehingga dianggap sebagai titik performa puncak yang optimal. Gambar 2, Evaluasi final menggunakan data uji dilakukan setelah mendapatkan *hyperparameter* terbaik yang telah divalidasi. Kinerja model dianalisis secara komprehensif melalui *Confusion Matrix, Classification Report*, dan ROC-AUC.



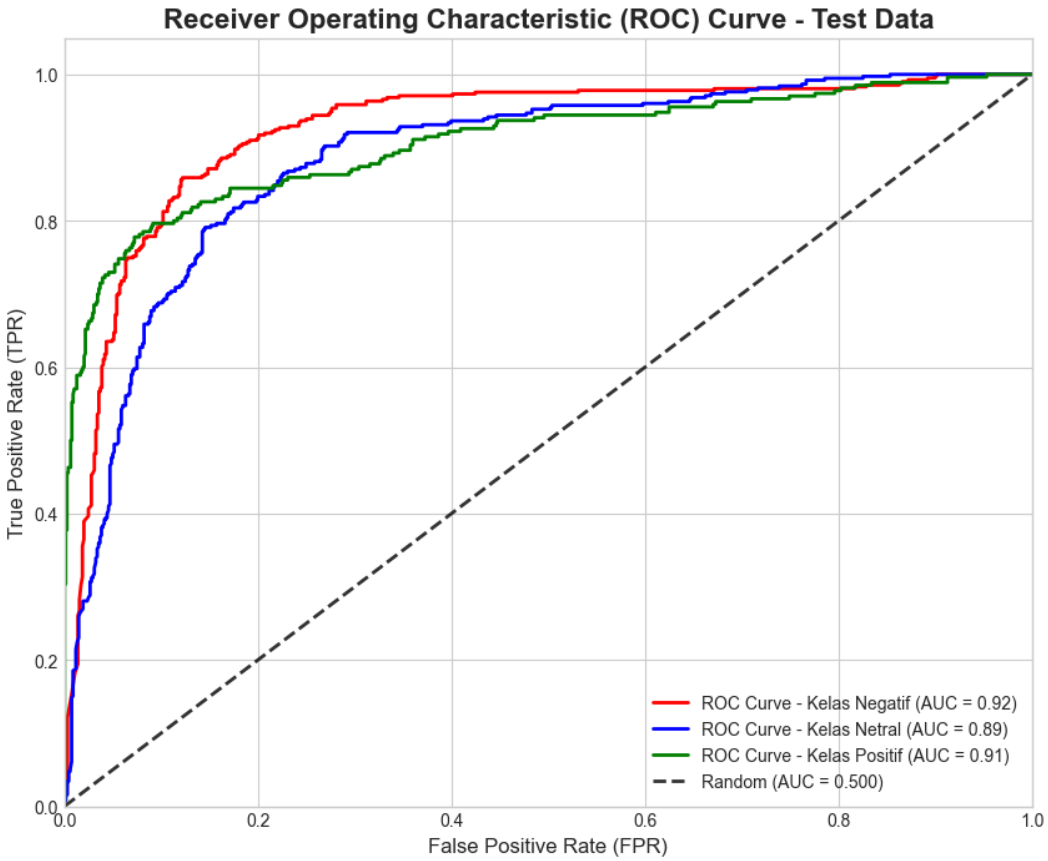
Gambar 2. Implementasi *Confusion Matrix*

Classification Report:				
	precision	recall	f1-score	support
Negatif	0.82	0.86	0.84	411
Netral	0.75	0.79	0.77	378
Positif	0.85	0.73	0.78	270
accuracy			0.80	1059
macro avg	0.81	0.79	0.80	1059
weighted avg	0.80	0.80	0.80	1059

Gambar 3. Hasil *Classification Report*

Analisis performa setiap kelas yang dapat dilihat pada gambar 2 dan 3 menunjukkan kemampuan model yang bervariasi. Kelas negatif merupakan performa paling kuat dengan *recall* tertinggi, menandakan kemampuan model yang sangat baik dalam mengidentifikasi sentimen negatif secara komprehensif. Sebaliknya, kelas Positif unggul dalam hal *precision*, yang berarti prediksinya sangat dapat diandalkan ketika menyatakan suatu sentimen adalah positif, namun memiliki *recall* terendah yang mengindikasikan model masih melewatkan sebagian dari total

sentimen positif yang sebenarnya. Adapun netral menunjukkan performa yang seimbang dengan *f1-score* sebesar 0.77.



Gambar 4. Hasil ROC-AUC

Hasil menunjukkan bahwa semua kelas memiliki nilai AUC di atas 89%, yang menunjukkan bahwa model memiliki daya diskriminasi yang sangat tinggi. Model paling andal dalam membedakan sentimen negatif dari sentimen lainnya, diikuti oleh sentimen positif. Performa yang tinggi ini menegaskan bahwa model IndoBERT yang telah dioptimalkan sangat efektif untuk tugas klasifikasi sentimen pada sentimen program makan siang gratis.

SIMPULAN

Penelitian ini telah berhasil mengevaluasi kinerja model IndoBERT dalam menganalisis sentimen terhadap program makan siang gratis di media sosial X. Melalui serangkaian optimasi sistematis, ditemukan konfigurasi *hyperparameter* terbaik yaitu pembagian data 68:12:20, *batch size* 8, pelatihan selama 7 *epoch*, *learning rate* 3e−5, dan *epsilon* 1e−9. Dengan konfigurasi tersebut, model final menunjukkan hasil performa yang kuat dan komprehensif, mencakup akurasi 80%, macro precision 0.81, macro recall 0.79, dan macro f1-score 0.80, serta kemampuan diskriminasi kelas yang sangat tinggi dengan rata-rata AUC 91%. Hasil ini membuktikan bahwa IndoBERT merupakan kerangka kerja yang andal dan efektif untuk analisis sentimen pada wacana berbahasa Indonesia.

Meskipun demikian, penelitian ini memiliki beberapa keterbatasan yaitu pada keterbatasan sumber data hanya terbatas pada platform media sosial X, sehingga hasilnya mungkin belum sepenuhnya merepresentasikan opini publik dari platform lain atau masyarakat luas.

DAFTAR PUSTAKA

[1] B. Gunawan and B. Ratmono, *Medsos Di Antara Dua Kutub*. 2021. Accessed: Apr. 28, 2025. [Online]. Available: https://www.google.co.id/books/edition/MEDSOS_di_Antara_Dua_Kutub/Vdk2EAAAQBAJ?hl=id&gbpv=1&dq=media+sosial&printsec=frontcover

[2] R. A. Salam and T. Akib, "Analisis Wacana Kritis terhadap Sarkasme dalam Twitter Sejak Bulan September-November 2023," Pendidikan, 2024. [Online]. Available: https://e-journal.my.id/onoma

[3] S. Kemp, "DIGITAL 2025," 2025.

[4] D. Bhisetya Rarasati, A. Pramana Thenata, A. Salsabila Arief, and C. Author, "Indonesian Society's Sentiment Analysis Against the COVID-19 Booster Vaccine," *Jurnal Nasional Teknik Elektro Dan Teknologi Informasi*, 2023.

[5] A. Sitanggang, Y. Umaidah, Y. Umaidah, R. I. Adam, and R. I. Adam, "Analisis Sentimen Masyarakat Terhadap Program Makan Siang Gratis Pada Media Sosial X Menggunakan Algoritma Naïve Bayes," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 12, no. 3, Mar. 2024, doi: 10.23960/jitet.v12i3.4902.

[6] L. Zhang, "Features extraction based on Naive Bayes algorithm and TF-IDF for news classification," *PLoS One*, vol. 20, no. 7 July, Jul. 2025, doi: 10.1371/journal.pone.0327347.

[7] Jamaaluddin and I. Sulistyowati, *Buku Ajar Kecerdasan Buatan (Artificial Intelligence)*. UMSIDA Press, 2021.

[8] A. S. Sulaeman, A. Sujjada, and I. L. Kharisma, "Penerapan Algoritma Cerdas Bidirectional Encoder Refresentations From Transformers Dalam Menganalisis Opini Publik Terhadap Produk Yang Mengalami Boikot," *Jurnal Inovtek Polbeng - Seri Informatika*, vol. 9, no. 1, 2024.

[9] L. Pradana, "Analisis Sentimen Masyarakat Media Sosial Twitter Terhadap Kinerja Penjabat Gubernur Dki Jakarta Menggunakan Model Indobert Program Studi Matematika," 2024.

[10] H. Imaduddin, F. Y. A'la, and Y. S. Nugroho, "Sentiment Analysis in Indonesian Healthcare Applications using IndoBERT Approach." [Online]. Available: www.ijacsa.thesai.org

[11] L. Geni, E. Yulianti, and D. I. Sensuse, "Sentiment Analysis of Tweets Before the 2024 Elections in Indonesia Using Bert Language Models," *Jurnal Ilmiah Teknik Elektro Komputer dan Informatika*, vol. 9, no. 3, pp. 746–757, Mar. 2023, doi: 10.26555/jiteki.v9i3.26490.

[12] F. Valerian, M. Syarief, and D. Fatah, "Klasifikasi Tingkat Obesitas Menggunakan Metode Gbm Dan Confusion Matrix," *Jurnal Mahasiswa Teknik Informatika*, vol. 9, no. 2, 2025.

[13] S. Sathyanarayanan and B. Tantri, "Confusion Matrix-Based Performance Evaluation Metrics," *African Journal of Biomedical Research*, pp. 4023–4031, Nov. 2024, doi: 10.53555/AJBR.v27i4S.4345.

[14] J. M. Rožanec and D. M. Mladeni´c, "WaAUROCC: measuring how steep the ROC curve is," 2024.

[15] A. Cahyana and E. Redy Susanto, "Penerapan Algoritma XGBoost untuk Prediksi Diabetes: Analisis Confusion Matrix dan ROC Curve," *Fountain of Informatics Journal*, vol. 10, no. 1, pp. 2548–5113, 2025, doi: 10.21111/fij.v10i1.14311.