PERBANDINGAN MODEL MULTILINGUAL INFOXLM DAN MDEBERTA DALAM ANALISIS SENTIMEN PROGRAM MBG DI MEDIA SOSIAL

TUGAS AKHIR
SKRIPSI

ARI SATRIO MURDOKO ANDJALMO
41822010025

PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2026

lib.mercubuana.ac.id

PERBANDINGAN MODEL MULTILINGUAL INFOXLM DAN MDEBERTA DALAM ANALISIS SENTIMEN PROGRAM MBG DI MEDIA SOSIAL

TUGAS AKHIR
SKRIPSI

Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh Gelar Sarjana

ARI SATRIO MURDOKO ANDJALMO
41822010025

PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2026

lib.mercubuana.ac.id

HALAMAN PERNYATAAN KARYA SENDIRI

Saya yang bertanda tangan di bawah ini:

Nama	:	ARI SATRIO MURDOKO ANDJALMO
NIM	:	41822010025
Program Studi	:	SISTEM INFORMASI

Menyatakan dengan sebenar-benarnya bahwa Tugas Akhir berjudul:
" PERBANDINGAN MODEL MULTILINGUAL INFOXLM DAN MDEBERTA DALAM ANALISIS SENTIMEN PROGRAM MBG DI MEDIA SOSIAL "
adalah hasil karya saya sendiri, tidak mengandung unsur plagiarisme, pelanggaran hak cipta, atau konten ilegal dalam bentuk apapun dan tidak melanggar hukum atau hak pihak manapun.
Apabila di kemudian hari ditemukan pelanggaran terhadap pernyataan ini, saya bersedia menanggung seluruh konsekuensi hukum dan membebaskan Universitas Mercu Buana dari segala bentuk tuntutan hukum dan saya siap mendapatkan sanksi akademis yang berlaku di Universitas Mercu Buana.
Demikian pernyataan ini saya buat dengan sebenar-benarnya untuk digunakan sebagaimana mestinya.

Jakarta, 20 Agustus 2026



ARI SATRIO MURDOKO ANDJALMO

UNIVERSITAS
MERCU BUANA

lib.mercubuana.ac.id

UNIVERSITAS
MERCU BUANA
AKREDITASI UNGGUL

PERNYATAAN SIMILARITY CHECK
/SIMILARITY CHECK STATEMENT

Saya yang bertanda tangan dibawah ini menyatakan, bahwa karya ilmiah yang ditulis oleh
/The undersigned, hereby declare that the scientific work written by

Nama /Name	: ARI SATRIO MURDOKO ANDJALMO
NIM /Student ID Number	: 41822010025
Program Studi /Study Program	: Sistem Informasi

Dengan Judul Tugas Akhir
/The title:

"PERBANDINGAN MODEL MULTILINGUAL INFOXLM DAN MDEBERTA DALAM ANALISIS SENTIMEN PROGRAM MBG DI MEDIA SOSIAL"

telah dilakukan pengecekan *similarity* dengan sistem Turnitin pada tanggal:
/Similarity checks have been carried out with the Turnitin system on the date:

18 Agustus 2026

dengan nilai persentase sebesar :
/with a percentage value of:

21%

dinyatakan memenuhi standar sesuai dengan ketentuan berlaku di **Fakultas Ilmu Komputer** Universitas Mercu Buana. */declared to meet standards in accordance with applicable regulations at the Faculty of Computer Science, Universitas Mercu Buana.*

File hasil cek similarity turnitin:
/Turnitin similarity report file

https://drive.google.com/file/d/169Yz8e0OgOy_5mGkWX-nZrhfc_jwiGzg/view?usp=drive_link

Jakarta, 18 Agustus 2026
Admin Turnitin Fasilkom UMB

Agung Prawoto

Agung Prawoto, S.Kom., B.Sc
NIK : 322970503

Fakultas Ilmu Komputer
KAMPUS MENARA BHAKTI
Jl. Raya Meruya Selatan No. 1 Kembangan, Jakarta Barat 11650
Telp. 021-5840816 (Hunting), Psw : 5700 Fax. 021-5840813
http://www.mercubuana.ac.id, e-mail : fasilkom@mercubuana.ac.id

lib.mercubuana.ac.id

HALAMAN PENGESAHAN

Laporan Skripsi ini diajukan oleh:

Nama	:	Ari Satrio Murdoko Andjalmo
NIM	:	41822010025
Program Studi	:	Sistem Informasi
Judul Laporan Skripsi	:	PERBANDINGAN MODEL MULTILINGUAL INFOXLM DAN MDEBERTA DALAM ANALISIS SENTIMEN PROGRAM MBG DI MEDIA SOSIAL

Telah berhasil dipertahankan pada sidang di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Strata 1 pada Program Studi Sistem Informasi, Fakultas Ilmu Komputer Universitas Mercu Buana.

Disahkan oleh:

Pembimbing,

Wawan Gunawan, S.Kom., M.T., M.Kom.
NIDN/NUPTK: 0424108104

UNIVERSITAS MERCU BUANA

Jakarta, 18 Agustus 2026

Mengetahui,

Dekan Fakultas Ilmu Komputer	Ketua Program Studi Sistem Informasi
Dr. Bambang Jokonowo, S.Si., M.T.I.	Wawan Gunawan, S.Kom., M.T., M.Kom.
NIDN: 0320037002	NIDN: 0424108104

lib.mercubuana.ac.id

KATA PENGANTAR

Puji syukur saya panjatkan kepada Tuhan Yang Maha Esa, karena atas berkat dan rahmat-Nya, saya dapat menyelesaikan Laporan Skripsi ini. Penulisan Laporan Skripsi ini dilakukan dalam rangka memenuhi salah satu syarat untuk mencapai gelar Sarjana Komputer pada Fakultas Ilmu Komputer Universitas Mercu Buana. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan skripsi ini, sangatlah sulit bagi saya untuk menyelesaikan Laporan Skripsi ini. Oleh karena itu, saya mengucapkan terima kasih kepada:

1. Prof Dr. Andi Adriansyah, M.Eng selaku Rektor Universitas Mercu Buana;
2. Assoc. Prof. Dr. Bambang Jokonowo, S.Si.,M.T.I selaku Dekan Fakultas Ilmu Komputer;
3. Wawan Gunawan, S.Kom., M.T., M.Kom selaku Ketua Program Studi Sistem Informasi serta selaku Dosen Pembimbing yang telah menyediakan waktu, tenaga, dan pikiran untuk mengarahkan saya dalam penyusunan skripsi ini;

Akhir kata, saya berharap Tuhan Yang Maha Esa berkenan membalas segala kebaikan semua pihak yang telah membantu. Semoga Laporan Skripsi ini membawa manfaat bagi pengembangan ilmu.

Jakarta, 28 Juli 2026



Penulis
ARI SATRIO MURDOKO ANDJALMO

lib.mercubuana.ac.id

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI
TUGAS AKHIR DI REPOSITORI UMB

Sebagai sivitas akademik Universitas Mercu Buana, saya yang bertanda tangan di bawah ini:

Nama	: ARI SATRIO MURDOKO ANDJALMO
NIM	: 41822010025
Fakultas/Program Studi	: ILMU KOMPUTER / SISTEM INFORMASI
Judul Tugas Akhir	: PERBANDINGAN MODEL MULTILINGUAL INFOXLM DAN MDEBERTA DALAM ANALISIS SENTIMEN PROGRAM MBG DI MEDIA SOSIAL

Demi pengembangan ilmu pengetahuan, dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Non-Eksklusif (*Non-exclusive Royalty-Free Right*)** atas karya ilmiah saya yang berjudul di atas beserta perangkat yang ada (jika diperlukan).
Dengan Hak Bebas Royalti Non-Eksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan Tugas Akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.
Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 20 Agustus 2026
Yang menyatakan,



(ARI SATRIO MURDOKO ANDJALMO)

UNIVERSITAS MERCU BUANA

lib.mercubuana.ac.id

PERBANDINGAN MODEL MULTILINGUAL INFOXLM DAN MDEBERTA DALAM ANALISIS SENTIMEN PROGRAM MBG DI MEDIA SOSIAL

ARI SATRIO MURDOKO ANDJALMO

ABSTRAK

Program Makan Bergizi Gratis (MBG) merupakan kebijakan strategis pemerintah Indonesia yang bertujuan meningkatkan kualitas gizi masyarakat, khususnya anak usia sekolah, dan dalam implementasinya memunculkan beragam respons publik di media sosial terkait kualitas makanan yang disediakan. Penelitian ini bertujuan menganalisis sentimen publik terhadap kualitas makanan Program MBG menggunakan data media sosial berbahasa Indonesia serta membandingkan kinerja model multilingual InfoXLM dan mDeBERTa dalam klasifikasi sentimen. Data penelitian dikumpulkan dari platform YouTube dan X/Twitter, kemudian melalui tahapan pra-pemrosesan dan pelabelan otomatis berbasis Large Language Model (LLM) DeepSeek v4 Pro yang divalidasi melalui anotasi manual terhadap 100 sampel acak untuk membentuk data ground truth. Analisis sentimen dilakukan dengan pendekatan supervised learning melalui proses fine-tuning pada kedua model, dengan evaluasi performa menggunakan metrik akurasi, precision, recall, dan F1-score. Hasil penelitian menunjukkan bahwa opini publik didominasi oleh sentimen negatif sebesar 81,73%, diikuti sentimen netral 12,93% dan positif 5,35%, yang mengindikasikan tingginya ketidakpuasan masyarakat terhadap kualitas makanan Program MBG. Dari sisi perbandingan model, mDeBERTa terbukti lebih unggul dan stabil dengan rata-rata akurasi 88,04% dan macro F1-score 0,7682, dibandingkan InfoXLM yang memperoleh akurasi 86,09% dan macro F1-score 0,6599, sehingga mDeBERTa dapat menjadi acuan pemilihan model dalam analisis sentimen kebijakan publik berbasis data..

Kata Kunci : Makan Bergizi, InfoXLM, Analisis Sentimen, mDeBERTa,Media Sosial.

UNIVERSITAS MERCU BUANA

lib.mercubuana.ac.id

COMPARISON OF MULTILINGUAL INFOXLM AND MDEBERTA MODELS IN SENTIMENT ANALYSIS OF THE MBG PROGRAM ON SOCIAL MEDIA

ARI SATRIO MURDOKO ANDJALMO

ABSTRACT

The Free Nutritious Meal Program (Program Makan Bergizi Gratis, MBG) is a strategic public policy implemented by the Indonesian government to improve nutritional quality, particularly among school-aged children, and its implementation has generated diverse public responses on social media regarding food quality. This study aims to analyze public sentiment toward the food quality of the MBG program using Indonesian-language social media data and to compare the performance of the multilingual models InfoXLM and mDeBERTa in sentiment classification. The data are collected from YouTube and X/Twitter, followed by text preprocessing and automatic labeling using the DeepSeek v4 Pro Large Language Model, validated through manual annotation of 100 random samples to construct the ground truth data. The results show that public opinion is dominated by negative sentiment (81.73%), followed by neutral (12.93%) and positive (5.35%) sentiment, indicating strong public dissatisfaction with the food quality of the MBG program. In terms of model comparison, mDeBERTa outperforms InfoXLM with an average accuracy of 88.04% and a macro F1-score of 0.7682, compared to 86.09% and 0.6599 respectively, demonstrating that mDeBERTa is more effective and stable in handling informal and imbalanced Indonesian social media text.

Keywords: *Free Nutritious Meal Program, Sentiment Analysis, InfoXLM, mDeBERTa, Social Media.*

UNIVERSITAS
MERCU BUANA

lib.mercubuana.ac.id

DAFTAR ISI

HALAMAN SAMPUL 0
HALAMAN JUDUL ii
HALAMAN PERNYATAAN KARYA SENDIRI ii
HALAMAN SURAT KETERANGAN UJI TURNITIN iii
HALAMAN PENGESAHAN iv
KATA PENGANTAR v
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR DI REPOSITORI UMB vi
ABSTRAK vii
ABSTRACT viii
DAFTAR ISI viii
DAFTAR TABEL xi
DAFTAR GAMBAR xii
DAFTAR LAMPIRAN xiii
BAB I PENDAHULUAN 1
1.1 Latar Belakang 1
1.2 Perumusan Masalah 4
1.3 Tujuan Penellitian 5
1.4 Batasan Masalah 5
1.5 Sistematika Penulisan 5
BAB II TINJAUAN PUSTAKA 7
2.1 Teori Utama 7
2.1.1 BERT(Bidirectional encoder representations from transformers) 7
2.1.2 InfoXLM 8
2.1.3 mDeBERTa 9
2.2 Teori Pendukung 10
2.2.1 NLP(Natural Language Processing) 10
2.2.2 Makan Bergizi Gratis (MBG) 11
2.2.3 Analisis Sentimen 11
2.2.4 Confusion Matrix dan Metrik Evaluasi Model 12
2.2.5 Pelabelan Data Berbasis Large Language Model (LLM) 14
2.3 Penelitian Terdahulu 15
2.4 Gap Penelitian 34
BAB III METODE PENELITIAN 36
3.1 Pendekatan Penelitian 36
3.2 Desain Penelitian 36
3.3 Objek dan Subjek Penelitian 38
3.3.1 Objek penelitian 38

lib.mercubuana.ac.id

3.3.2 Subjek Penelitian 38
3.4 Instrumen Penelitian 38
3.5 Tahapan Penelitian 39
3.6 Teknik Pengumpulan Data 40
3.7 Teknik Analisis dan Validasi Model 40
BAB IV HASIL DAN PEMBAHASAN 42
4.1 Pengumpulan Data 42
4.1.1 X (Twitter) 42
4.1.2 Youtube 45
4.2 Pra-Pemrosesan data 47
4.2.1 Cleaning Data 47
4.2.2 Labeling Data 51
4.2.3 Normalisasi 55
4.3 Distribusi Sentimen Publik 59
4.4 Pembagian Dataset 60
4.5 Permodelan 60
4.5.1 Implementasi Model InfoXLM 62
4.5.2 Implementasi Model mDeBERTa 63
4.6 Evaluasi Model 65
BAB V KESIMPULAN DAN SARAN 70
5.1 Kesimpulan 70
5.2 Saran 71
DAFTAR PUSTAKA 73
LAMPIRAN 75

UNIVERSITAS
MERCU BUANA

lib.mercubuana.ac.id

DAFTAR TABEL

Tabel 2.1 Struktur Confusion Matrix 12
Tabel 2.2 Penelitian Terdahulu 15
Tabel 3.1 Sample Dataset 37
Tabel 3.2 Spesifikasi Hardware 38
Tabel 3.3 Versi Software 38
Tabel 4.1 Konfigurasi twitter crawler 42
Tabel 4.2 Tabel hasil crawling twitter 43
Tabel 4.3 Daftar video yang dipakai 45
Tabel 4.4 hasil dari scrapping youtube 46
Tabel 4.5 Kode untuk feature selection 47
Tabel 4.6 Header dataset sebelum dan setelah feature selection 48
Tabel 4.7 Tabel sebelum dan sesudah whitespace removal 49
Tabel 4.8 Tabel sebelum dan sesudah pembersihan link URL 50
Tabel 4.9 Hasil pembersihan mention serta hashtag 50
Tabel 4.10 Hasil pembersihan emoji 51
Tabel 4.11 Prompt yang digunakan untuk labeling pada AI 52
Tabel 4.12 Contoh data relevan dan tidak relevan 53
Tabel 4.13 Sample komparasi antara 53
Tabel 4.14 Tabel hasil komparasi 55
Tabel 4.15 Hasil dari case folding 56
Tabel 4.16 Sample mapping slang 56
Tabel 4.17 Hasil pembersihan ekspresi tawa 57
Tabel 4.18 Hasil Normalisasi 57
Tabel 4.19 Hasil Tokenisasi 58
Tabel 4.20 Hasil Fine-Tuning Model InfoXLM 62
Tabel 4.21 Hasil fine tuning model mDeBERTa 64
Tabel 4.22 Performa InfoXLM dan mDeBERTa pada kelas sentimen Positif 65
Tabel 4.23 Performa InfoXLM dan mDeBERTa pada kelas sentiment Netral 66
Tabel 4.24 Performa InfoXLM dan mDeBERTa pada kelas sentimen Negatif 67

UNIVERSITAS MERCU BUANA

lib.mercubuana.ac.id

DAFTAR GAMBAR

Gambar 2.1. Contextual Word Embeddings dalam BERT 7
Gambar 3.1 Desain Penelitian 37
Gambar 3.2 Tahapan Penelitian 40
Gambar 4.1 panjang token 61
Gambar 4.2 Training dan Validation loss dari mDeBERTa 67
Gambar 4.3 Training dan Validation loss dari Model InfoXLM 68
Gambar 4.4 Confusion matrix dari Model InfoXLM (kiri) & Model MdeBERTa (kanan) 69

UNIVERSITAS
MERCU BUANA

lib.mercubuana.ac.id

DAFTAR LAMPIRAN

Lampiran 1 Kartu Asistensi 75
Lampiran 2 CV 76
Lampiran 3 Bukti BNSP 77

UNIVERSITAS
MERCU BUANA

lib.mercubuana.ac.id