

**ANALISIS SENTIMEN PUBLIK TERHADAP PROGRAM
MAKAN BERGIZI GRATIS (MBG) DI PLATFORM X
MENGUNAKAN ALGORITMA INDOBERTWEET**



SKRIPSI

**VIVO HIZKIA IMANUEL
00000060115**

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2026**

**ANALISIS SENTIMEN PUBLIK TERHADAP PROGRAM
MAKAN BERGIZI GRATIS (MBG) DI PLATFORM X
MENGUNAKAN ALGORITMA INDOBERTWEET**



SKRIPSI

Diajukan sebagai salah satu syarat untuk memperoleh
Gelar Sarjana Komputer (S.Kom.)

**VIVO HIZKIA IMANUEL
00000060115**

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2026**

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Vivo Hizkia Imanuel

Nomor Induk Mahasiswa : 00000060115

Program Studi : Informatika

Skripsi dengan judul:

Analisis Sentimen Publik terhadap Program Makan Bergizi Gratis (MBG) di Platform X Menggunakan Algoritma INDOBERTweet

merupakan hasil karya saya sendiri bukan plagiat dari laporan karya tulis ilmiah yang ditulis oleh orang lain, dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar serta dicantumkan di Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/penyimpangan, baik dalam pelaksanaan maupun dalam penulisan laporan karya tulis ilmiah, saya bersedia menerima konsekuensi dinyatakan TIDAK LULUS untuk mata kuliah yang telah saya tempuh.

Tangerang, 8 Januari 2026



(Vivo Hizkia Imanuel)

UMM
UNIVERSITAS
MULTIMEDIA
NUSANTARA

HALAMAN PENGESAHAN

Skripsi dengan judul

ANALISIS SENTIMEN PUBLIK TERHADAP PROGRAM MAKAN BERGIZI GRATIS (MBG) DI PLATFORM X MENGGUNAKAN ALGORITMA INDOBERTWEET

oleh

Nama : Vivo Hizkia Imanuel
NIM : 00000060115
Program Studi : Informatika
Fakultas : Fakultas Teknik dan Informatika

Telah diujikan pada hari Kamis, 13 Januari 2026

Pukul 10.00 s/s 12.00 dan dinyatakan

LULUS

Dengan susunan penguji sebagai berikut

Ketua Sidang

(Wirawan Istiono, S.Kom., M.Kom.)

NIDN: 0313048304

Penguji

(Dr. Maria Irmira Prasetyowati, S.Kom., M.T.)

NIDN: 0725057201

Pembimbing

(Eunike Endariahna Surbakti, S.Kom., M.T.I.)

NIDN: 0322099401

Ketua Program Studi Informatika,

(Arya Wicaksana, S.Kom., M.Eng.Sc., OCA)

NIDN: 0315109103

**HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK
KEPENTINGAN AKADEMIS**

Yang bertanda tangan di bawah ini:

Nama : Vivo Hizkia Imanuel
NIM : 00000060115
Program Studi : Informatika
Jenjang : S1
Judul Karya Ilmiah : Analisis Sentimen Publik terhadap
Program Makan Bergizi Gratis
(MBG) di Platform X Menggunakan
Algoritma INDOBERTweet

Menyatakan dengan sesungguhnya bahwa saya bersedia (**pilih salah satu**):

- ☒ Saya bersedia memberikan izin sepenuhnya kepada Universitas Multimedia Nusantara untuk mempublikasikan hasil karya ilmiah saya ke dalam repositori Knowledge Center sehingga dapat diakses oleh Sivitas Akademika UMN/Publik. Saya menyatakan bahwa karya ilmiah yang saya buat tidak mengandung data yang bersifat konfidensial.
- ☐ Saya tidak bersedia mempublikasikan hasil karya ilmiah ini ke dalam repositori Knowledge Center, dikarenakan: dalam proses pengajuan publikasi ke jurnal/konferensi nasional/internasional (dibuktikan dengan *letter of acceptance*) **.
- ☐ Lainnya, pilih salah satu:
- Hanya dapat diakses secara internal Universitas Multimedia Nusantara
 - Embargo publikasi karya ilmiah dalam kurun waktu tiga tahun.

Tangerang, 8 Januari 2026

Yang menyatakan



Vivo Hizkia Imanuel

**Jika tidak bisa membuktikan LoA jurnal/HKI, saya bersedia mengizinkan penuh karya ilmiah saya untuk dipublikasikan ke KC UMN dan menjadi hak institusi UMN.

HALAMAN PERSEMBAHAN / MOTTO

”Di dalam Yesus Kristus tersembunyi segala harta karun hikmat dan pengetahuan Allah yang sejati, menegaskan bahwa Kristus adalah pusat dari kebenaran ilahi, dan menjadi peringatan agar tidak terpengaruh oleh ajaran-ajaran duniawi yang menyesatkan.”

Kolose 2:3



KATA PENGANTAR

Segala puji bagi Tuhan Yang Maha Esa atas segala limpahan karunia-Nya, yang memungkinkan penulis menyelesaikan Tugas Akhir ini. Penyusunan laporan ini dimaksudkan sebagai salah satu syarat akademik guna menyandang gelar Sarjana Komputer pada Program Studi Informatika, Universitas Multimedia Nusantara.

Selama proses penelitian dan penulisan, penulis banyak mendapatkan bantuan, arahan, serta inspirasi dari berbagai pihak yang sangat berarti. Atas segala kebaikan tersebut, penulis mengucapkan terima kasih kepada:

1. Bapak Dr. Ir. Andrey Andoko, M.Sc., selaku Rektor Universitas Multimedia Nusantara.
2. Bapak Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Dekan Fakultas Teknik dan Informatika Universitas Multimedia Nusantara.
3. Bapak Arya Wicaksana, S.Kom., M.Eng.Sc., OCA, selaku Ketua Program Studi Informatika Universitas Multimedia Nusantara.
4. Ibu Eunike Endariahna Surbakti, S.Kom., M.T.I., sebagai Pembimbing pertama yang telah memberikan bimbingan, arahan, dan motivasi atas terselesainya tugas akhir ini.
5. Terimakasih kepada Kevin Sorensen S. Kom, Joshua Ardin Putra S. Kom, Kevin Melky, Rasyid Alim Aulia S. Kom, Abidzar Awal, Hansen S. Kom, Agil Wira S. Kom, Farhan S. Kom yang setia membantu dan menemani saya dari awal kuliah hingga sekarang.
6. Terimakasih kepada sahabat dari "anak baik" dan "akhirnya bebas dari kesesakan" telah menemani hari-hari saya.
7. Saya mengucapkan terima kasih yang mendalam kepada orang tua saya atas pembiayaan pendidikan dari awal hingga akhir, dengan penghormatan khusus untuk almarhum Bapak.
8. Terimakasih kepada orang yang sudah mendukung saya baik secara verbal maupun non verbal

9. Terimakasih kepada sahabat dari "Core Graha" yang sudah membantu saya melewati banyak hal.
10. Terimakasih kepada sahabat saya "besok badmin" yang telah menemani sehari-hari.

Penulis menyadari bahwa tugas akhir ini masih banyak kekurangan, oleh karena itu yang diharapkan penelitian ini bisa bermanfaat dan menjadi referensi untuk pengembang.

Tangerang, 8 Januari 2026



Vivo Hizkia Imanuel



**ANALISIS SENTIMEN PUBLIK TERHADAP PROGRAM MAKAN
BERGIZI GRATIS (MBG) DI PLATFORM X MENGGUNAKAN
ALGORITMA INDOBERTWEET**

Vivo Hizkia Imanuel

ABSTRAK

Program Makan Bergizi Gratis (MBG) merupakan salah satu kebijakan prioritas pemerintah Indonesia yang memicu beragam respons masyarakat di media sosial, khususnya *platform X*. Penelitian ini bertujuan untuk menganalisis sentimen publik terhadap program tersebut dengan menerapkan model *IndoBERTweet*, sebuah varian BERT yang dioptimalkan untuk teks Twitter berbahasa Indonesia. Data penelitian dikumpulkan melalui teknik *crawling*, menghasilkan 10.498 *tweet*. Yang dibagi menjadi *training* (70%), *validation* (15%), dan *testing* (15%). Hasil penelitian menunjukkan bahwa kombinasi strategi *pseudo-labeling* dengan prapemrosesan lengkap memberikan performa model terbaik pada konfigurasi *learning rate* 2×10^{-5} , menghasilkan nilai *Accuracy* 97,67%, *Precision* 97,76%, *Recall* 97,67%, dan *F1-Score* 97,70%. Distribusi sentimen dalam dataset *pseudo-labeling* menunjukkan sentimen positif sebesar 48,14%, sentimen netral sebesar 32,74%, dan sentimen negatif sebesar 19,12%. Temuan ini mengindikasikan bahwa *pseudo-labeling* sebagai strategi peningkatan data berlabel secara signifikan meningkatkan performa model *IndoBERTweet*, dan model terbukti efektif dalam mengklasifikasikan sentimen pada teks informal berbahasa Indonesia dengan tingkat akurasi yang tinggi.

Kata kunci: Analisis Sentimen, *IndoBERTweet*, Program MBG, *Pseudo-labeling*, Media Sosial X.

U M N
U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

***PUBLIC SENTIMENT ANALYSIS OF THE FREE NUTRITIOUS MEAL
(MBG) PROGRAM ON PLATFORM X USING THE INDOBERTWEET
ALGORITHM***

Vivo Hizkia Imanuel

ABSTRACT

The Free Nutritious Meal (MBG) program is one of the priority policies of the Indonesian government that has triggered diverse public responses on social media, particularly on platform X. This study aims to analyze public sentiment towards the program by applying the IndoBERTweet model, a BERT variant optimized for Indonesian Twitter text. Research data were collected via crawling techniques, yielding 10,498 tweets which were divided into training (70%), validation (15%), and testing (15%). The results showed that the combination of pseudo-labeling strategy with complete preprocessing yielded the best model performance at a learning rate configuration of 2×10^{-5} , resulting in an Accuracy of 97.67%, Precision of 97.76%, Recall of 97.67%, and F1-Score of 97.70%. Sentiment distribution in the pseudo-labeling dataset showed positive sentiment at 48.14%, neutral sentiment at 32.74%, and negative sentiment at 19.12%. These findings indicate that pseudo-labeling as a labeled data augmentation strategy significantly improved the performance of the IndoBERTweet model, and the model proved effective in classifying sentiment in informal Indonesian text with high accuracy.

Keywords Sentiment Analysis, IndoBERTweet, MBG Program, Pseudo-labeling, Social Media X.

UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR ISI

HALAMAN JUDUL	i
PERNYATAAN TIDAK MELAKUKAN PLAGIAT	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH	iv
HALAMAN PERSEMBAHAN/MOTO	v
KATA PENGANTAR	vi
ABSTRAK	viii
ABSTRACT	ix
DAFTAR ISI	x
DAFTAR TABEL	xii
DAFTAR GAMBAR	xiii
DAFTAR KODE	xiv
DAFTAR RUMUS	xv
DAFTAR LAMPIRAN	xvi
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	3
1.3 Batasan Permasalahan	3
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	4
1.6 Sistematika Penulisan	4
BAB 2 LANDASAN TEORI	5
2.1 Analisis Sentimen	5
2.2 Platform X (sebelumnya Twitter)	5
2.3 <i>Tweet Harvest</i>	5
2.4 <i>Text Processing</i>	6
2.5 <i>Heuristic Labeling</i>	7
2.6 <i>Pseudo Labeling</i>	7
2.7 <i>Class Weight</i>	8
2.8 <i>Natural Language Processing</i>	8
2.9 <i>Model Transformer</i>	9
2.10 Model BERT	9
2.10.1 Mekanisme <i>Output</i> dan <i>Fine-tuning</i>	11
2.11 IndoBERT	12
2.12 IndoBERTweet	12
2.13 Evaluasi	13
BAB 3 METODOLOGI PENELITIAN	14
3.1 Gambaran Umum Penelitian	14
3.2 Studi Literatur	15
3.3 Arsitektur	15
3.4 Pengumpulan Data	16
3.4.1 Perancangan Model	17
3.4.2 <i>Preprocessing</i>	18
3.4.3 <i>Model Training</i>	22
3.5 Evaluasi	23
BAB 4 HASIL DAN DISKUSI	24
4.1 Karakteristik dan Klasifikasi Sentimen Data	24
4.2 Spesifikasi Perangkat	26

4.3	Pengumpulan Data	26
4.4	Perancangan Sistem	27
4.5	<i>Preprocessing</i>	29
4.6	Perbandingan Performa Model terhadap Skenario <i>Labeling</i> dan Metode <i>Preprocessing</i>	34
4.6.1	Skenario 1: <i>Heuristic Labeling</i> dengan <i>Preprocessing Default</i>	35
4.6.2	Skenario 2: <i>Heuristic Labeling</i> dengan <i>Preprocessing Paper</i>	36
4.6.3	Skenario 3: <i>Pseudo Labeling</i> dengan <i>Preprocessing Default</i>	38
4.6.4	Skenario 4: <i>Pseudo Labeling</i> dengan <i>Preprocessing Paper</i> .	39
4.6.5	Perbandingan Keempat Skenario	40
4.6.6	<i>Heuristic Labeling</i>	41
4.6.7	<i>Pseudo Labeling</i>	45
4.7	Konfigurasi <i>Epoch & Learning Rate</i>	49
4.8	Analisis Komparatif: <i>Heuristic</i> dan <i>Pseudo Labeling</i>	50
4.9	Dampak <i>Hyperparameter</i> terhadap Performa Model	51
BAB 5	SIMPULAN DAN SARAN	52
5.1	Kesimpulan	52
5.2	Saran	53
DAFTAR PUSTAKA		54



DAFTAR TABEL

Tabel 3.1	Parameter Konfigurasi Model	16
Tabel 4.1	Definisi dan Karakteristik Kelas Sentimen Program MBG .	25
Tabel 4.2	Rangkuman Skenario <i>Testing</i>	35
Tabel 4.3	Hasil Validasi Skenario 1: <i>Heuristic Labeling</i> + <i>Preprocessing Default</i>	36
Tabel 4.4	Hasil Validasi Skenario 2: <i>Heuristic Labeling</i> + <i>Preprocessing Paper</i>	37
Tabel 4.5	Hasil Validasi Skenario 3: <i>Pseudo Labeling</i> + <i>Preprocessing Default</i>	38
Tabel 4.6	Hasil Validasi Skenario 4: <i>Pseudo Labeling</i> + <i>Preprocessing Paper</i>	39
Tabel 4.7	Perbandingan Performa Terbaik Keempat Skenario <i>Testing</i>	40
Tabel 4.8	Perbandingan Performa Model Berdasarkan Metode <i>Preprocessing</i>	42
Tabel 4.9	Perbandingan Performa Model Berdasarkan Metode <i>Preprocessing - Pseudo Labeling</i>	46
Tabel 4.10	<i>Best Epoch</i> dan <i>Test F1-Score</i> untuk 12 eksperimen (2 <i>Labeling</i> × 2 <i>Preprocessing</i> × 3 <i>LR</i>)	49



DAFTAR GAMBAR

Gambar 2.1	Arsitektur model <i>Transformer</i> . Sumber: [19]	9
Gambar 2.2	Visualisasi representasi <i>input</i> pada BERT yang terdiri dari <i>Token</i> , <i>Segment</i> , dan <i>Position Embeddings</i> . Sumber: [20]	10
Gambar 2.3	Ilustrasi proses <i>pre-training</i> BERT menggunakan <i>Masked Language Modeling</i> (MLM) dan <i>Next Sentence Prediction</i> (NSP). Sumber: [20]	11
Gambar 3.1	Diagram Alur Metodologi Penelitian	14
Gambar 3.2	<i>Flowchart</i> Pengumpulan Data	17
Gambar 3.3	<i>Flowchart Text Preprocessing</i>	18
Gambar 3.4	Contoh kata dari <i>case folding</i>	19
Gambar 3.5	Contoh untuk menghapus URL, #, <i>mention</i> , dan emoji	19
Gambar 3.6	Contoh tokenisasi	20
Gambar 3.7	Contoh <i>stopword</i>	20
Gambar 3.8	Contoh menormalisasikan kata baku	21
Gambar 3.9	Contoh kata berulang	21
Gambar 3.10	<i>Flowchart</i> Pelatihan Model	22
Gambar 3.11	<i>Flowchart Testing</i>	23
Gambar 4.1	Perbandingan metrik validasi berdasarkan <i>epoch</i> dan <i>learning rate</i> pada Skenario 1	36
Gambar 4.2	Perbandingan metrik validasi berdasarkan <i>epoch</i> dan <i>learning rate</i> pada Skenario 2	37
Gambar 4.3	Perbandingan metrik validasi berdasarkan <i>epoch</i> dan <i>learning rate</i> untuk Skenario 3	38
Gambar 4.4	Perbandingan metrik validasi berdasarkan <i>epoch</i> dan <i>learning rate</i> untuk Skenario 4	39
Gambar 4.5	Perbandingan <i>Confusion Matrix Training</i>	43
Gambar 4.6	Perbandingan <i>Confusion Matrix Validation</i>	44
Gambar 4.7	Perbandingan <i>Confusion Matrix Testing</i>	44
Gambar 4.8	Perbandingan <i>Confusion Matrix Training Set - Pseudo Default vs Pseudo Paper</i>	47
Gambar 4.9	Perbandingan <i>Confusion Matrix Validation Set - Pseudo Default vs Pseudo Paper</i>	48
Gambar 4.10	Perbandingan <i>Confusion Matrix Testing Set - Pseudo Default vs Pseudo Paper</i>	49

UNIVERSITAS
MULTIMEDIA
NUSANTARA

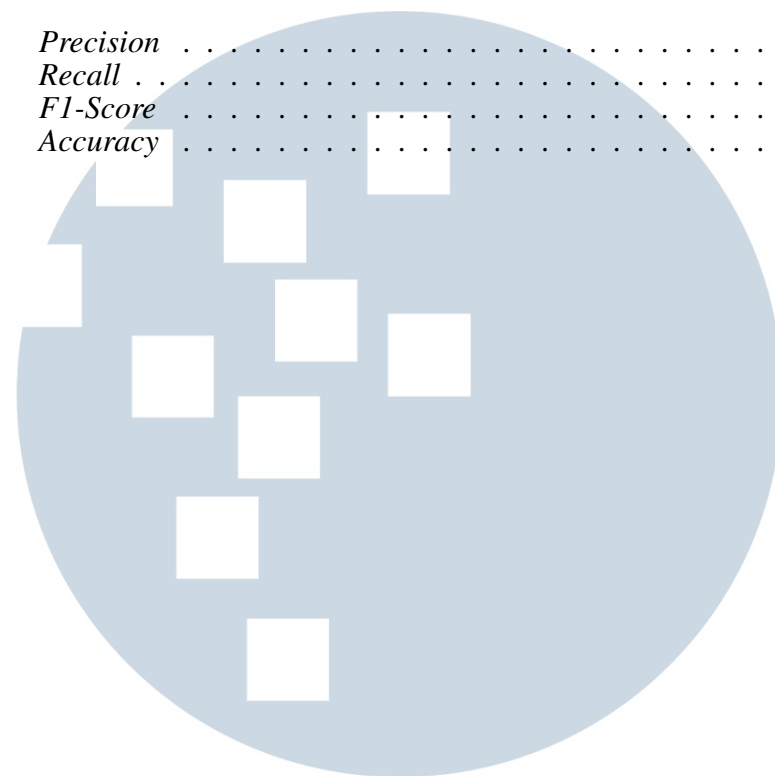
DAFTAR KODE

Kode 4.1	Inisialisasi Token Autentikasi Twitter	26
Kode 4.2	<i>Harvesting Tweet</i> dengan <i>Tweet-Harvest</i>	27
Kode 4.3	Penggabungan dan Deduplikasi File CSV	28
Kode 4.4	Proses <i>Case Folding</i>	29
Kode 4.5	Pembersihan Teks Tahap 2: Menghapus <i>URL</i> , Angka, dan Karakter Khusus	30
Kode 4.6	Penghapusan Mention dan Hashtag	31
Kode 4.7	Penghapusan Stopword dari Teks	32
Kode 4.8	Normalisasi Teks Berdasarkan Kamus	33
Kode 4.9	Tokenisasi Teks Menggunakan BERT	34



DAFTAR RUMUS

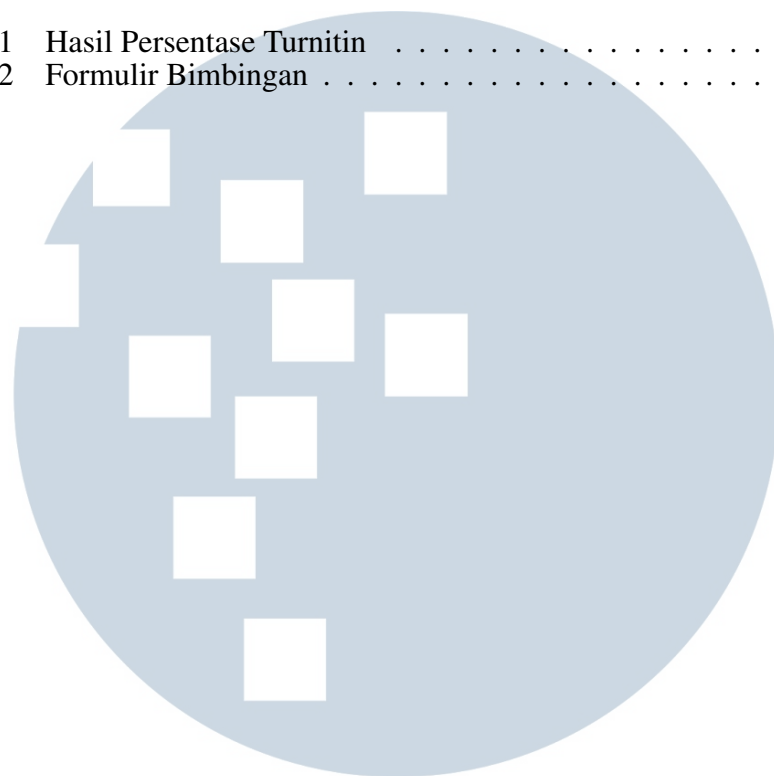
Rumus 2.1	<i>Precision</i>	13
Rumus 2.2	<i>Recall</i>	13
Rumus 2.3	<i>F1-Score</i>	13
Rumus 2.4	<i>Accuracy</i>	13



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR LAMPIRAN

Lampiran 1	Hasil Persentase Turnitin	57
Lampiran 2	Formulir Bimbingan	58



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA