

ANALISIS SENTIMEN TERHADAP PROGRAM MAKAN BERGIZI GRATIS (MBG) MENGGUNAKAN MODEL INDOBERT DAN BIGRU

SKRIPSI

Oleh:

**REINHARD PRASETYA
NIM. 221112453**



**PROGRAM STUDI S-1 TEKNIK INFORMATIKA
FAKULTAS INFORMATIKA
UNIVERSITAS MIKROSKIL
MEDAN
2026**

**SENTIMENT ANALYSIS OF INDONESIA'S MAKAN BERGIZI
GRATIS (MBG) PROGRAM USING THE INDOBERT
AND BIGRU MODELS**

UNDERGRADUATE THESIS

By:

**REINHARD PRASETYA
ID NUMBER. 221112453**



**UNDERGRADUATE PROGRAM OF COMPUTER SCIENCE
FACULTY OF INFORMATICS
UNIVERSITAS MIKROSKIL
MEDAN
2026**

LEMBARAN PENGESAHAN

**ANALISIS SENTIMEN TERHADAP PROGRAM MAKAN BERGIZI
GRATIS (MBG) MENGGUNAKAN MODEL INDOBERT DAN BIGRU**

SKRIPSI

Diajukan untuk Melengkapi Persyaratan Guna
Mendapatkan Gelar Sarjana
Program Studi S-1 Teknik Informatika

Oleh:

REINHARD PRASETYA
NIM. 221112453

Disetujui Oleh:

Dosen Pembimbing I,


Sunaryo Winardi, S.Kom., M.T.

Dosen Pembimbing II,


Darwin, S.Kom., M.Kom.

Medan, 10 Agustus 2026
Diketahui dan Disahkan Oleh:

Ketua Program Studi
S-1 Teknik Informatika


Carles Juliandy, S.Kom., M.Kom.

HALAMAN PERNYATAAN

Saya yang membuat pernyataan ini adalah mahasiswa Program Studi S-1 Teknik Informatika Universitas Mikroskil Medan dengan identitas mahasiswa sebagai berikut:

Nama : Reinhard Prasetya

NIM : 221112453

Saya telah melaksanakan penelitian dan penulisan Skripsi dengan judul dan tempat penelitian sebagai berikut:

Judul Skripsi : Analisis Sentimen terhadap Program Makan Bergizi Gratis (MBG) Menggunakan Model IndoBERT dan BiGRU

Tempat Penelitian : -

Alamat Tempat Penelitian : -

No. Telp. Tempat Penelitian : -

Sehubungan dengan Skripsi tersebut, dengan ini saya menyatakan dengan sebenar-benarnya bahwa penelitian dan penulisan Skripsi tersebut merupakan hasil karya saya sendiri (tidak menyuruh orang lain yang mengerjakannya) dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar. Bila di kemudian hari ternyata terbukti bahwa bukan saya yang mengerjakannya (membuatnya), maka saya bersedia dikenakan sanksi yang telah ditetapkan oleh Universitas Mikroskil Medan, yakni pencabutan ijazah yang telah saya terima dan ijazah tersebut dinyatakan tidak sah.

Selain itu, demi pengembangan ilmu pengetahuan, saya menyetujui untuk memberikan kepada Universitas Mikroskil Medan Hak Bebas Royalti Non-eksklusif (Non-exclusive Royalty Free Right) atas Skripsi saya beserta perangkat yang ada (jika diperlukan). Dengan hak ini, Universitas Mikroskil Medan berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan Skripsi saya, secara keseluruhan atau hanya sebagian atau hanya ringkasannya saja dalam bentuk format tercetak dan/atau elektronik, selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik hak cipta. Menyatakan juga bahwa saya akan mempertahankan hak eksklusif saya untuk menggunakan seluruh atau sebagian isi Skripsi saya guna pengembangan karya di masa depan, misalnya dalam bentuk artikel, buku, ataupun perangkat lunak/sistem informasi.

Demikian pernyataan ini saya perbuat dengan sungguh-sungguh, dalam keadaan sadar dan tanpa ada tekanan dari pihak manapun.

Medan, 6 Juli 2026

Saya yang membuat pernyataan,



Reinhard Prasetya

ANALISIS SENTIMEN TERHADAP PROGRAM MAKAN BERGIZI GRATIS (MBG) MENGGUNAKAN MODEL INDOBERT DAN BIGRU

Abstrak

Program Makan Bergizi Gratis (MBG) menimbulkan beragam opini masyarakat di media sosial Twitter/X, sehingga diperlukan metode analisis sentimen otomatis. Penelitian ini bertujuan mengembangkan sistem analisis sentimen berbasis website yang mengklasifikasikan opini masyarakat terhadap Program MBG ke dalam tiga kelas sentimen (positif, netral, negatif) menggunakan arsitektur hybrid IndoBERT-BiGRU. Data latih berjumlah 2.970 tweet dengan distribusi kelas tidak seimbang (negatif 1.921, positif 926, netral 123), sehingga diterapkan oversampling hingga menjadi 5.763 data dengan distribusi seimbang. Pencarian parameter fine-tuning dilakukan pada 27 kombinasi skenario, dengan dan tanpa oversampling. Hasil tertinggi diperoleh pada kombinasi oversampling dengan batch size 32, learning rate 3e-5, dan epoch 5 (checkpoint epoch keempat), dengan accuracy, precision weighted, recall weighted, dan F1-score weighted validasi masing-masing 0,9352, 0,9346, 0,9352, dan 0,9347. Pada data uji, keempat metrik tersebut sebesar 0,9191, 0,9183, 0,9191, dan 0,9184, dengan selisih F1-score 0,1579 poin dibandingkan tanpa oversampling.

Kata kunci: Analisis Sentimen; IndoBERT-BiGRU; Media Sosial; Oversampling; Program Makan Bergizi Gratis

Abstract

Program Makan Bergizi Gratis (MBG) has led to various public opinions on the social media platform Twitter/X, requiring an automatic sentiment analysis method. This research aims to develop a website-based sentiment analysis system that classifies public opinion on the MBG Program into three sentiment classes (positive, neutral, negative) using a hybrid IndoBERT-BiGRU architecture. The training data consisted of 2,970 tweets with an imbalanced class distribution (negative 1,921, positive 926, neutral 123), so oversampling was applied, resulting in 5,763 balanced data. A search for the best fine-tuning parameters was carried out across 27 scenario combinations, with and without oversampling. The best results were obtained using oversampling, with a batch size of 32, a learning rate of 3e-5, and 5 epochs (checkpoint at epoch four), producing validation accuracy, weighted precision, weighted recall, and weighted F1-score of 0.9352, 0.9346, 0.9352, and 0.9347. On the test data, the four metrics were 0.9191, 0.9183, 0.9191, and 0.9184, with an F1-score difference of 0.1579 points compared to the model without oversampling.

Keywords: IndoBERT-BiGRU; Makan Bergizi Gratis Program; Oversampling; Sentiment Analysis; Social Media

KATA PENGANTAR

Puji dan syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa atas segala rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan Tugas Akhir ini dengan baik, yang berjudul “ANALISIS SENTIMEN TERHADAP PROGRAM MAKAN BERGIZI GRATIS (MBG) MENGGUNAKAN MODEL INDOBERT DAN BIGRU”.

Penulisan skripsi ini dilakukan dalam rangka memenuhi salah satu syarat untuk mencapai gelar Sarjana Komputer pada Program Studi Teknik Informatika, Fakultas Informatika, Universitas Mikroskil.

Penulis menyadari bahwa penulisan Skripsi ini tidak mungkin diselesaikan tanpa adanya bantuan dan dukungan dari berbagai pihak yang terlibat. Untuk itu, pada kesempatan ini penulis ingin menyampaikan rasa terima kasih yang sebesar-besarnya kepada:

1. Bapak Sunaryo Winardi, S.Kom., M.T., selaku Dosen Pembimbing I, dan juga selaku Dekan Fakultas Informatika Universitas Mikroskil Medan.
2. Bapak Darwin, S.Kom., M.Kom., selaku Dosen Pembimbing II.
3. Bapak Hardy, S.Kom., M.Sc., Ph.D., selaku Rektor Universitas Mikroskil Medan.
4. Bapak Carles Juliandy, S.Kom., M.Kom., selaku Ketua Program Studi S-1 Teknik Informatika, Fakultas Informatika Universitas Mikroskil Medan.
5. Kepada orang tua yang senantiasa memberikan dukungan secara doa, motivasi, dan juga material untuk penulis selama mengikuti pendidikan hingga selesainya Tugas Akhir.
6. Kepada sahabat dan teman-teman dekat yang telah memberikan perhatian, dukungan, serta semangat selama proses penyusunan Tugas Akhir ini.

Penulis menyadari bahwa Tugas Akhir ini masih memiliki keterbatasan dan kekurangan. Oleh karena itu, penulis sangat mengharapkan kritik dan saran yang bersifat membangun demi penyempurnaan di masa mendatang. Akhir kata, semoga Tugas Akhir ini dapat memberikan manfaat bagi pembaca dan perkembangan ilmu pengetahuan. Terima kasih.

Medan, 6 Juli 2026

Penulis,

Reinhard Prasetya

DAFTAR ISI

ABSTRAK.....	i
KATA PENGANTAR	ii
DAFTAR ISI	iii
DAFTAR GAMBAR	v
DAFTAR TABEL.....	vi
DAFTAR LAMPIRAN	viii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan.....	3
1.4 Manfaat.....	3
1.5 Ruang Lingkup	4
BAB II KAJIAN LITERATUR.....	5
2.1 <i>Natural Language Processing</i> (NLP).....	5
2.2 <i>Text Mining</i>	7
2.3 Analisis Sentimen.....	9
2.4 <i>Lexicon-Based</i>	10
2.5 <i>Oversampling</i> Berbasis Augmentasi Teks.....	12
2.6 IndoBERT	13
2.7 BiGRU	19
2.8 <i>Hybrid Model</i>	22
2.9 <i>Confusion Matrix</i>	24
2.10 X	26
2.11 Makan Bergizi Gratis	27
BAB III TAHAPAN PELAKSANAAN	29
3.1 Analisis Kebutuhan	30

3.1.1 Analisis Kebutuhan Fungsional.....	30
3.1.2 Analisis Kebutuhan Non-Fungsional	34
3.2 Perancangan Sistem.....	35
3.2.1 Perancangan Tampilan	35
3.2.2 Perancangan Proses	40
BAB IV HASIL DAN PEMBAHASAN	87
4.1 Hasil.....	87
4.1.1 Tampilan Halaman Beranda.....	87
4.1.2 Tampilan Halaman Analisis Sentimen	88
4.1.3 Tampilan Halaman Hasil Analisis Sentimen.....	89
4.1.4 Tampilan Halaman <i>Fine-Tuning</i> Model.....	89
4.1.5 Tampilan Halaman Evaluasi Performa Model	91
4.2 Pembahasan.....	93
4.2.1 Hasil Pelatihan Model tanpa <i>Oversampling</i>	94
4.2.2 Hasil Pelatihan Model menggunakan <i>Oversampling</i>	98
4.2.3 Perbandingan Hasil Pelatihan Model	101
4.2.4 Pengujian Model pada Data <i>Tweet</i> Nyata	103
BAB V PENUTUP	107
5.1 Kesimpulan.....	107
5.2 Saran.....	107
DAFTAR PUSTAKA	108
DAFTAR RIWAYAT HIDUP.....	116

DAFTAR GAMBAR

Gambar 2. 1 Arsitektur umum IndoBERT berbasis encoder BERT [51].....	14
Gambar 2. 2 Struktur BiGRU [56]	21
Gambar 2. 3 Pengguna Aktif Twitter 2017-2023 [76]	26
Gambar 3. 1 <i>Use Case Diagram</i>	38
Gambar 3. 2 Halaman Utama	44
Gambar 3. 3 Halaman Analisis Sentimen.....	45
Gambar 3. 4 Halaman Hasil Analisis	45
Gambar 3. 5 Halaman <i>Fine-tuning</i> Model.....	46
Gambar 3. 6 Halaman Evaluasi Performa Model.....	47
Gambar 3. 7 <i>Flowchart</i> Proses Sistem	48
Gambar 3. 8 Struktur <i>Dataset</i> Mentah (<i>Raw Dataset</i>) dari Kaggle.....	49
Gambar 3. 9 <i>Flowchart Preprocessing</i>	49
Gambar 3. 10 Perbandingan Jumlah Data sebelum dan sesudah <i>Data Cleansing</i>	61
Gambar 3. 11 Distribusi Kelas Sentimen setelah <i>Split Dataset</i>	65
Gambar 4. 1 Tampilan Halaman Beranda	95
Gambar 4. 2 Tampilan Halaman Analisis Sentimen	96
Gambar 4. 3 Tampilan Halaman Hasil Analisis Sentimen.....	97
Gambar 4. 4 Tampilan Halaman <i>Fine-Tuning</i> Model Bagian Atas	98
Gambar 4. 5 Tampilan Halaman <i>Fine-Tuning</i> Model Bagian Bawah.....	98
Gambar 4. 6 Tampilan <i>Loading Bar</i> Saat Proses <i>Training</i> Model Berjalan	99
Gambar 4. 7 Tampilan Halaman Evaluasi Performa Model Bagian Atas.....	100
Gambar 4. 8 Tampilan Halaman Evaluasi Performa Model Bagian Bawah	100
Gambar 4. 9 Distribusi Kelas Sentimen tanpa <i>Oversampling</i>	102
Gambar 4. 10 Grafik <i>Training Loss</i> , <i>Validation Loss</i> , <i>Training Accuracy</i> , dan <i>Validation Accuracy</i> Model dengan Nilai Tertinggi tanpa <i>Oversampling</i>	104
Gambar 4. 11 <i>Confusion matrix</i> Hasil Pengujian Model tanpa <i>Oversampling</i>	105
Gambar 4. 12 Grafik <i>Training Loss</i> , <i>Validation Loss</i> , <i>Training Accuracy</i> , dan <i>Validation Accuracy</i> Model dengan Nilai Tertinggi menggunakan <i>Oversampling</i>	108
Gambar 4. 13 <i>Confusion matrix</i> Hasil Pengujian Model menggunakan <i>Oversampling</i> ...	109

DAFTAR TABEL

Tabel 3. 1 Definisi Aktor	37
Tabel 3. 2 Skenario <i>Use Case</i> : Melakukan analisis sentimen	37
Tabel 3. 3 Skenario <i>Use Case</i> : Melakukan <i>Fine-tuning</i> Model	38
Tabel 3. 4 Kebutuhan Non-fungsional Sistem	40
Tabel 3. 5 Tahapan <i>Casefolding</i>	48
Tabel 3. 6 Tahapan <i>Remove Non-Latin</i>	49
Tabel 3. 7 Penanganan <i>HTML Entities</i>	50
Tabel 3. 8 Tahapan <i>Remove URL</i>	52
Tabel 3. 9 Tahapan <i>Remove Hashtag</i>	53
Tabel 3. 10 Tahapan <i>Remove Mention</i>	54
Tabel 3. 11 Tahapan <i>Remove Numbers</i>	56
Tabel 3. 12 Tahapan <i>Remove Punctuation</i>	56
Tabel 3. 13 Tahapan <i>Replace Slang</i>	57
Tabel 3. 14 Tahapan <i>Normalize Extra</i>	57
Tabel 3. 15 Tahapan <i>Replace Word Elongation</i>	58
Tabel 3. 16 Tahapan <i>Clean Whitespace</i>	58
Tabel 3. 17 Perhitungan Skor pada Kalimat dan Penentuan Label Sentimen	60
Tabel 3. 18 Distribusi Label	61
Tabel 3. 19 Contoh Hasil Augmentasi	62
Tabel 3. 20 Distribusi Label setelah <i>Oversampling</i>	62
Tabel 3. 21 Hasil Tokenisasi	63
Tabel 3. 22 Penambahan token khusus IndoBERT	63
Tabel 3. 23 Token diubah menjadi <i>Numeric ID</i>	64
Tabel 3. 24 <i>Padding</i>	64
Tabel 3. 25 <i>Token Embedding</i>	65
Tabel 3. 26 <i>Position Embedding</i>	66
Tabel 3. 27 <i>Segment Embedding</i>	66
Tabel 3. 28 Hasil Penjumlahan <i>Embedding</i>	67
Tabel 3. 29 Mean dan Standar Deviasi per Token	70
Tabel 3. 30 Parameter γ (gamma) dan β (beta)	70
Tabel 3. 31 Hasil <i>Embedding</i> setelah <i>Layer Normalization</i>	72
Tabel 3. 32 <i>Query, Key, Value (Layer 1)</i>	73

Tabel 3. 33 <i>Attention Score Head 0</i>	75
Tabel 3. 34 <i>Attention Weight Head 0 (Setelah Softmax)</i>	77
Tabel 3. 35 <i>Attention Weight Head 6</i>	77
Tabel 3. 36 <i>Attention Weight Head 11</i>	77
Tabel 3. 37 <i>Output Self-Attention (Setelah Concatenation, Residual Connection, dan Layer normalization)</i>	79
Tabel 3. 38 <i>Output Intermediate FFN (768 → 3.072, Setelah GELU)</i>	79
Tabel 3. 39 <i>Output Layer 1 Encoder Transformer (Setelah FFN, Residual Connection, dan Layer Normalization)</i>	80
Tabel 3. 40 Perubahan Vektor Token “makan” di Setiap <i>Layer</i>	81
Tabel 3. 41 <i>Output Akhir IndoBERT (Layer 12)</i>	83
Tabel 3. 42 <i>Gate Values Forward GRU</i>	85
Tabel 3. 43 <i>Output Backward GRU</i>	86
Tabel 3. 44 Hasil <i>Concatenation</i> seluruh Token	88
Tabel 3. 45 Alur Transformasi Nilai Token "makan" Dimensi ke-0.....	90
Tabel 4. 1 Hasil Pelatihan Model tanpa <i>Oversampling</i>	102
Tabel 4. 2 Ringkasan Hasil <i>Training</i> Tertinggi tanpa <i>Oversampling</i>	104
Tabel 4. 3 Hasil Pelatihan Model menggunakan <i>Oversampling</i>	106
Tabel 4. 4 Ringkasan Hasil <i>Training</i> Tertinggi menggunakan <i>Oversampling</i>	107
Tabel 4. 5 Selisih Performa Model dengan Nilai Tertinggi berdasarkan Metrik Evaluasi	110
Tabel 4. 6 Perbandingan Hasil <i>Confusion Matrix</i> Model Tanpa <i>Oversampling</i> dengan Model Menggunakan <i>Oversampling</i>	111
Tabel 4. 7 Hasil Prediksi 10 Data <i>Tweet</i> Nyata.....	112

DAFTAR LAMPIRAN

Lampiran 1 <i>Link</i>	119
Lampiran 2 Tabel.....	119

