

KLASIFIKASI SENTIMEN BERBASIS GPT DAN PEMODELAN TOPIK LDA TERHADAP OPINI PUBLIK PADA PLATFORM X MENGENAI ISU PROGRAM MAKAN BERGIZI GRATIS

MERIZA IMMANUELA VIRGIE



**PROGRAM STUDI SARJANA STATISTIKA DAN SAINS DATA
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2026**



@Hak cipta milik IPB University

IPB University

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber ;
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

PERNYATAAN MENGENAI SKRIPSI DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa skripsi dengan judul “Klasifikasi Sentimen Berbasis GPT dan Pemodelan Topik LDA terhadap Opini Publik pada Platform X mengenai Isu Program Makan Bergizi Gratis” adalah karya saya dengan arahan dari dosen pembimbing dan belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir skripsi ini.

Dalam penyusunan karya ini, penulis menggunakan bantuan kecerdasan buatan berupa ChatGPT sebagai bagian dari penelitian, selain itu digunakan pula Claude dan Gemini untuk perbaikan kode program, penyusunan kerangka penulisan skripsi, dan penyuntingan tata bahasa yang digunakan. Setelah menggunakan alat/layanan tersebut, penulis meninjau dan menyunting konten sesuai kebutuhan serta bertanggung jawab penuh atas isi karya tugas akhir ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Institut Pertanian Bogor.

Bogor, Agustus 2026

Meriza Immanuela Virgie
G1401221076

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

ABSTRAK

MERIZA IMMANUELA VIRGIE. Klasifikasi Sentimen Berbasis GPT dan Pemodelan Topik LDA terhadap Opini Publik pada Platform X mengenai Isu Program Makan Bergizi Gratis. Dibimbing oleh CICI SUHAENI dan MOHAMMAD MASJKUR.

Penelitian ini bertujuan untuk menganalisis persepsi publik terhadap Program Makan Bergizi Gratis (MBG) di Indonesia melalui pemodelan topik dan analisis sentimen terhadap data media sosial X. Sebanyak 7.004 tweet dikumpulkan dan diproses melalui beberapa tahapan *preprocessing*, yaitu *filtering*, pembersihan teks, normalisasi, tokenisasi, *stopword removal*, dan *stemming*. Pemodelan topik dilakukan menggunakan metode *Latent Dirichlet Allocation* (LDA), sedangkan klasifikasi sentimen dilakukan menggunakan model *Generative Pre-trained Transformer* (GPT). Hasil penelitian menunjukkan bahwa model LDA optimal menghasilkan empat topik utama, yaitu (1) Pelaksanaan dan Kontroversi Program MBG, (2) Kebijakan Pangan dan Dinamika Politik, (3) Pengembangan Sumber Daya Manusia dan Prioritas Anggaran, serta (4) Isu Pendukung dan Diskursus Kontekstual. Di antara keempat topik tersebut, pembahasan mengenai pelaksanaan program dan berbagai kontroversi yang menyertainya merupakan topik yang paling dominan. Model GPT menunjukkan kinerja klasifikasi sentimen yang baik, meskipun masih memiliki keterbatasan dalam mengklasifikasikan sentimen netral. Model ini memperoleh nilai akurasi sebesar 84,01%, *balanced accuracy* sebesar 72,28%, *weighted F1-score* sebesar 84,94%, dan *macro F1-score* sebesar 68,92%. Analisis sentimen berdasarkan waktu dan topik menunjukkan bahwa sentimen negatif secara konsisten mendominasi diskusi publik pada platform X selama periode pengamatan maupun pada seluruh topik yang diidentifikasi. Temuan tersebut mengindikasikan bahwa persepsi publik di X terhadap Program MBG masih didominasi oleh kritik, kekhawatiran, dan ketidakpuasan, terutama terkait pelaksanaan program serta kebijakan-kebijakan yang mendukungnya.

Kata kunci: Analisis Sentimen, *Large Language Models* (LLM), *Latent Dirichlet Allocation* (LDA), Makan Bergizi Gratis (MBG), Pemodelan Topik

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber ;
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.

2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

ABSTRACT

MERIZA IMMANUELA VIRGIE. GPT-Based Sentiment Classification and LDA Topic Modeling of Public Opinion on the X Platform Regarding the Free Nutritious Meals Program. Supervised by CICI SUHAENI and MOHAMMAD MASJKUR.

This study aims to analyze public perceptions of the Free Nutritious Meals Program (MBG) in Indonesia through topic modeling and sentiment analysis of social media data from X. A total of 7,004 tweets were collected and processed through several preprocessing steps, namely filtering, text cleaning, normalization, tokenization, stopword removal, and stemming. Topic modeling was performed using the Latent Dirichlet Allocation (LDA) method, while sentiment classification was performed using the Generative Pre-trained Transformer (GPT) model. The results of the study show that the optimal LDA model yielded four main topics, namely (1) Implementation and Controversies of the MBG Program, (2) Food Policy and Political Dynamics, (3) Human Resource Development and Budget Priorities, and (4) Supporting Issues and Contextual Discourse. Among these four topics, discussions regarding the program's implementation and the various controversies surrounding it were the most dominant. The GPT model demonstrated good sentiment classification performance, although it still has limitations in classifying neutral sentiments. This model achieved an accuracy of 84.01%, a balanced accuracy of 72.28%, a weighted F1-score of 84.94%, and a macro F1-score of 68.92%. Sentiment analysis based on time and topic shows that negative sentiment consistently dominated public discussions during the observation period and across all identified topics. These findings indicate that public perception on X regarding the MBG Program is still dominated by criticism, concerns, and dissatisfaction, particularly regarding the program's implementation and the policies supporting it.

Keywords: Large Language Models (LLM), Latent Dirichlet Allocation (LDA), Makan Bergizi Gratis (MBG), sentiment analysis, topic modeling

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber ;
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



@Hak cipta milik IPB University

IPB University



- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

© Hak Cipta milik IPB, tahun 2026
Hak Cipta dilindungi Undang-Undang

Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan atau menyebutkan sumbernya. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah, dan pengutipan tersebut tidak merugikan kepentingan IPB.

Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apa pun tanpa izin IPB.

KLASIFIKASI SENTIMEN BERBASIS GPT DAN PEMODELAN TOPIK LDA TERHADAP OPINI PUBLIK PADA PLATFORM X MENGENAI ISU PROGRAM MAKAN BERGIZI GRATIS

MERIZA IMMANUELA VIRGIE

Skripsi
sebagai salah satu syarat untuk memperoleh gelar
Sarjana pada
Program Studi Statistika dan Sains Data

**PROGRAM STUDI SARJANA STATISTIKA DAN SAINS DATA
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2026**



@Hak cipta milik IPB University

IPB University

Tim Penguji pada Ujian Skripsi:

1 Dr. Agus Mohamad Soleh, S.Si., M.T



- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Judul Skripsi : Klasifikasi Sentimen Berbasis GPT dan Pemodelan Topik LDA terhadap Opini Publik pada Platform X mengenai Isu Program Makan Bergizi Gratis

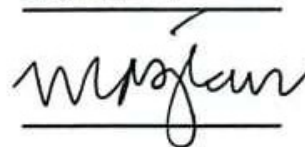
Nama : Meriza Immanuela Virgie

NIM : G1401221076

Disetujui oleh

Pembimbing 1:
Cici Suhaeni, S.Si, M.Si, Ph.D

Pembimbing 2:
Ir. Mohammad Masjkur M.S

Diketahui oleh

Ketua Program Studi:
Dr. Bagus Sartono, S.Si., M.Si.
NIP 19780411 200501 1 002

Tanggal Ujian:
27 Juli 2026

Tanggal Lulus:

PRAKATA

Puji dan syukur penulis panjatkan kepada Tuhan Yesus Kristus atas segala karunia-Nya sehingga karya ilmiah ini berhasil diselesaikan. Tema yang dipilih dalam penelitian yang dilaksanakan sejak bulan Februari 2026 sampai bulan Juni 2026 ini ialah *text analytics* untuk program Makan Bergizi Gratis, dengan judul “Klasifikasi Sentimen Berbasis GPT dan Pemodelan Topik LDA terhadap Opini Publik pada Platform X mengenai Isu Program Makan Bergizi Gratis”.

Penyelesaian karya ilmiah ini tidak terlepas dari bantuan, dukungan, serta kontribusi berharga dari berbagai pihak yang berperan penting selama proses penelitian berlangsung. Oleh karena itu, penulis menyampaikan rasa terima kasih yang sebesar-besarnya kepada:

1. Orang tua penulis, Bapak Amos Rizal dan Ibu Meriaty H. Purba, adik kandung penulis, Benmanuel Rizal, serta seluruh keluarga besar penulis. Dukungan, doa, serta kasih sayang yang diberikan menjadi sumber kekuatan utama dalam penyelesaian karya ilmiah ini.
2. Dosen pembimbing penulis, Ibu Cici Suhaeni, S.Si, M.Si, Ph.D dan Bapak Ir. Mohammad Masjkur M.S yang telah membimbing, mengarahkan, dan mendukung penuh penulis dalam pelaksanaan, penyusunan, dan penyelesaian penelitian ini.
3. Dosen penguji, Bapak Dr. Agus Mohamad Soleh, S.Si., M.T atas saran, masukan, serta kritik yang sangat membangun dan bermanfaat untuk perbaikan dan penyempurnaan penelitian ini
4. Dosen moderator kolokium serta seminar hasil, Bapak La Ode Abdul Rahman, S.Si., M.Si. dan Bapak Gerry Alfa Dito, S.Si., M.Si., atas *feedback* yang mendorong penulis untuk meningkatkan kualitas penelitian ini.
5. Seluruh dosen dan tenaga kependidikan Program Studi Statistika dan Sains Data. Dedikasi mereka dalam berbagi ilmu pengetahuan serta bantuan yang diberikan selama perjalanan akademik dan proses penyusunan karya ilmiah ini sangat penulis hargai.
6. Seluruh sahabat-sahabat terbaik penulis, terlebih khusus Nindy, Azanti, dan Anggre yang senantiasa menemani dan mendukung perjalanan penulis selama menjalani perkuliahan di Program Studi Statistika dan Sains Data
7. Kepada BPH SAM 2024 dan 2025, UluluSquad, QQ, Roy, info next badmin, dan Ulya atas kebersamaan, dukungan, serta doa-doa yang diberikan.
8. Marinestic Statistika 59, GSB 2024 dan 2025, Departemen *Sport and Art Management*, serta seluruh pihak yang selalu memberikan warna dan pengalaman yang berharga

Akhir kata, semoga karya ilmiah ini bermanfaat bagi pihak yang membutuhkan serta dapat berkontribusi bagi kemajuan ilmu pengetahuan.

Bogor, Agustus 2026

Meriza Immanuela Virgie



@Hak cipta milik IPB University

IPB University



IPB University
— Bogor Indonesia —

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber ;
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Perpustakaan IPB University

DAFTAR ISI

DAFTAR TABEL	xi
DAFTAR GAMBAR	xi
DAFTAR LAMPIRAN	xi
I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Tujuan	2
1.3 Ruang Lingkup	3
II TINJAUAN PUSTAKA	4
2.1 Program Makan Bergizi Gratis (MBG)	4
2.2 <i>Text Preprocessing</i>	4
2.3 <i>Topic Modeling</i>	5
2.4 Klasifikasi Sentimen	7
III METODE	10
3.1 Data	10
3.2 Prosedur Analisis Data	10
IV HASIL DAN PEMBAHASAN	15
4.1 Pengumpulan dan Penyiapan Data	15
4.2 Hasil <i>Text Preprocessing</i> dan Eksplorasi Data	16
4.3 Pemodelan Topik dengan LDA	19
4.4 Klasifikasi Sentimen dengan LLM	25
4.5 Analisis Distribusi Sentimen	28
V SIMPULAN DAN SARAN	32
5.1 Simpulan	32
5.2 Saran	32
DAFTAR PUSTAKA	34
LAMPIRAN	37
RIWAYAT HIDUP	40

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber ;
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

DAFTAR TABEL

1	Contoh data yang terkumpul	15
2	Contoh <i>tweet</i> yang dieliminasi pada tahap <i>filtering</i>	16
3	Contoh kata yang mengalami <i>custom normalization</i>	17
4	Daftar contoh <i>custom stopwords</i>	18
5	Hasil <i>Text Preprocessing</i>	18
6	Metrik Evaluasi tiap Jumlah Topik	20
7	10 kata teratas tiap topik	22
8	Nilai Metrik Evaluasi Hasil <i>Labeling</i> Sentimen	26

DAFTAR GAMBAR

1	Diagram Alir Penelitian	10
2	Eksplorasi data (a) jumlah <i>tweet</i> per kuartal dan (b) distribusi panjang kata per <i>tweet</i>	19
3	<i>Word cloud</i> hasil <i>text preprocessing</i>	19
4	<i>Intertopic distance map</i> untuk (a) 3 topik, (b) 4 topik, (c) 5 topik, (d) 6 topik, dan (e) 7 topik	21
5	<i>Top 20</i> kata teratas untuk (a) topik 1, (b) topik 2, (c) topik 3, dan (d) topik 4	24
6	<i>Confussion Matrix</i> Perbandingan <i>Labeling</i> Sentimen Manual dan AI	27
7	Persentase masing-masing sentimen per kuartal	29
8	Distribusi sentimen pada topik 1, topik 2, topik 3, dan topik 4	30

DAFTAR LAMPIRAN

1	Lampiran 1 Prompt klasifikasi sentimen dengan GPT-5.4- <i>mini</i>	38
2	Lampiran 2 Contoh hasil pelabelan sentimen manual	39



@Hak cipta milik IPB University

IPB University

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.