



ANALISIS SENTIMEN PENGGUNA X TERHADAP PELAKSANAAN PROGRAM MBG MENGGUNAKAN METODE MULTINOMIAL NAIVE BAYES

JONATHAN HIZKIA BURJU SIMANJUNTAK



**PROGRAM STUDI SARJANA STATISTIKA DAN SAINS DATA
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2025**



@Hak cipta milik 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 “Analisis Sentimen Pengguna X Terhadap Pelaksanaan Program MBG Menggunakan Metode Multinomial Naive Bayes” 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.

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

Bogor, Juli 2025

Jonathan Hizkia Burju Simanjuntak
G1401211104



ABSTRAK

JONATHAN HIZKIA BURJU SIMANJUNTAK. Analisis Sentimen Pengguna X Terhadap Pelaksanaan Program MBG Menggunakan Metode Multinomial Naive Bayes. Dibimbing oleh SACHNAZ DESTA OKTARINA dan ANANG KURNIA.

Program Makan Bergizi Gratis (MBG) merupakan inisiatif pemerintah untuk menyediakan makanan bergizi secara gratis kepada pelajar, balita, ibu hamil, dan menyusui, namun efektivitasnya memerlukan evaluasi respons masyarakat. Penelitian ini bertujuan menggunakan algoritma *Multinomial Naive Bayes* untuk menganalisis sentimen komentar pengguna platform X terhadap pelaksanaan MBG. Algoritma MNB digunakan karena kemampuannya dalam menangani matriks fitur *sparse* serta mengolah data bervolume besar secara cepat, akurat, dan transparan. Data dikumpulkan pada periode awal pelaksanaan program (periode *baseline*) dan setelah program berjalan beberapa bulan (periode *endline*). Sebanyak 6875 komentar yang memuat istilah “MBG” dikumpulkan pada bulan Januari 2025 (periode *baseline*) dan bulan April 2025 (periode *endline*). Komentar tersebut diberikan label positif, netral, atau negatif melalui *hybrid AI-human annotation*, serta diproses meliputi pembersihan teks, tokenisasi, normalisasi, *stopword removal*, dan *stemming*. Fitur diekstraksi dengan menggunakan metode *term frequency (TF)*. Model MNB dengan *back-translation* menghasilkan *weighted F1-score* tertinggi sebesar 0,79. Model tersebut menunjukkan kinerja unggul dalam mendeteksi sentimen positif ($F1 = 0,84$) dan negatif ($F1 = 0,83$), walaupun klasifikasi netral masih terbatas ($F1 = 0,58$). Analisis temporal mengungkap pergeseran dominasi sentimen negatif (55 %) menjadi positif (58 %) pada tiga bulan setelah implementasi program. Temuan ini menegaskan efektivitas pendekatan MNB untuk pemantauan sentimen skala besar.

Kata kunci: analisis sentimen, evaluasi program, kecerdasan buatan, multinomial naive Bayes, program makan bergizi gratis.

@Hak cipta milik IPB University

ABSTRACT

JONATHAN HIZKIA BURJU SIMANJUNTAK. Sentiment Analysis of X Users on Free Nutritious Meal Program Implementation using Multinomial Naive Bayes. Supervised by SACHNAZ DESTA OKTARINA and ANANG KURNIA.

The Free Nutritious Meal Program (MBG) is a government initiative to provide free nutritious meals to students, toddlers, pregnant, and lactating women, yet its effectiveness requires evaluation of public response. This study aims to use the Multinomial Naive Bayes algorithm to analyze the sentiment of X user comments on MBG implementation. The MNB algorithm is used for its ability to handle sparse feature matrices and process large volumes of data quickly, accurately, and transparently. Data are collected during the initial period of program implementation (baseline period) and after the program has been running for several months (endline period). A total of 6,875 comments containing the term MBG were collected in Januari 2025 (baseline period) and April 2025 (endline period). Comments were given labels as positive, neutral, or negative through hybrid AI-human annotation and processed via text cleaning, tokenization, normalization, stopword removal, and stemming. Features were extracted using *term frequency* (TF). The back-translation MNB model achieved a highest weighted F1-score of 0,79. The model demonstrate strong performance in detecting positive ($F1 = 0,84$) and negative sentiments ($F1 = 0,83$), although neutral classification remained limited ($F1 = 0,58$). Temporal analysis revealed a shift from predominantly negative sentiment (55%) to predominantly positive sentiment (58%) after three months of program implementation. These findings confirm the effectiveness of the MNB approach for large-scale sentiment monitoring.

Keywords: artificial intelligence, free nutritious meal program, multinomial naive Bayes, program evaluation, sentiment analysis.



@Hak cipta milik 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 2025¹
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.



ANALISIS SENTIMEN PENGGUNA X TERHADAP PELAKSANAAN PROGRAM MBG MENGGUNAKAN METODE MULTINOMIAL NAIVE BAYES

JONATHAN HIZKIA BURJU SIMANJUNTAK

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
2025**



@Hak cipta milik 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 : Analisis Sentimen Pengguna X Terhadap Pelaksanaan Program
MBG Menggunakan Metode Multinomial Naive Bayes

Nama : Jonathan Hizkia Burju Simanjuntak
NIM : G1401211104

Disetujui oleh

Pembimbing 1:
Sachnaz Desta Oktarina, S.Stat., M.Agr., Ph.D.

Pembimbing 2:
Prof. Dr. Anang Kurnia S.Si., M.Si.



Diketahui oleh

Ketua Program Studi:
Dr. Bagus Sartono, S.Si., M.Si
NIP 197804112005011002



Tanggal Ujian:
(17 Juli 2025)

Tanggal Lulus:
(tanggal penandatanganan oleh Dekan
Fakultas/Sekolah ...)



- 

PRAKATA

Puji dan syukur penulis panjatkan kepada Tuhan yang Maha Esa atas segala karunia-Nya sehingga karya ilmiah ini berhasil diselesaikan. Tema yang dipilih dalam penelitian yang dilaksanakan sejak bulan Januari 2025 sampai bulan Juli 2025 ini ialah analisis sentimen, dengan judul “Analisis Sentimen Pengguna X Terhadap Pelaksanaan Program MBG Menggunakan Metode Multinomial Naive Bayes”. Pada proses pembuatan karya ilmiah ini, penulis mendapatkan doa, dukungan, dan bimbingan dari berbagai pihak. Oleh karena itu, penulis ingin memberikan ucapan terima kasih setinggi-tingginya kepada:

1. Bapak Rizky Anthony Bonardo Simanjuntak dan Ibu Vera Susana Simbolon sebagai orang tua penulis yang selalu memberikan doa, dukungan, dan motivasi dalam setiap langkah kehidupan penulis.
2. Ibu Sachnaz Desta Oktarina, S.Stat., M.Agr., Ph.D. dan Prof. Dr. Anang Kurnia S.Si., M.Si. selaku dosen pembimbing yang memberikan bimbingan dan arahan kepada penulis dalam menyelesaikan penulisan karya ilmiah.
3. Dr. Bagus Sartono, S.Si., M.Si selaku dosen penguji yang telah memberikan banyak saran dan masukan kepada penulis dalam menyelesaikan penulisan karya ilmiah ini.
4. Seluruh dosen dan staf Program Studi Statistika dan Sains Data, SSMI IPB yang membantu penulis selama perkuliahan sehingga dapat menyelesaikan studi dan administrasi sampai kelulusan.
5. Rebecca Zessica Anastasya Sihombing sebagai *partner* penulis yang telah senantiasa kebersamai serta memberikan doa, dukungan moral, dan berbagai bantuan kepada penulis dalam menyelesaikan penulisan karya ilmiah.
6. Gilang Dwi Laksana dan Muhammad Nafiz sebagai teman penulis yang telah kebersamai dan memberikan dukungan selama penulisan karya ilmiah ini.
7. Teman-teman dan sahabat Statistika IPB 58, PMK IPB, serta IDC IPB yang telah kebersamai dan memberikan dukungan selama proses perkuliahan.
8. Seluruh pihak yang tidak dapat disebutkan satu persatu yang sudah memberikan dukungan dan bantuan selama proses perkuliahan.

Semoga karya ilmiah ini bermanfaat bagi pihak yang membutuhkan dan bagi kemajuan ilmu pengetahuan.

Bogor, Agustus 2025

Jonathan Hizkia Burju Simanjuntak



@Hak cipta milik 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.



DAFTAR ISI

DAFTAR TABEL	ix
DAFTAR GAMBAR	ix
DAFTAR LAMPIRAN	ix
PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan	3
1.4 Manfaat	3
TINJAUAN PUSTAKA	4
2.1 Analisis Sentimen	4
2.2 <i>Multinomial Naive Bayes</i>	4
2.3 <i>Term Frequency (TF)</i>	5
2.4 <i>Large Language Models (LLM)</i>	6
2.5 Metode Penanganan Data Tidak Seimbang	7
2.6 Metode Evaluasi Model	7
III METODE	10
3.1 Data	10
3.2 Prosedur Kerja	10
IV HASIL DAN PEMBAHASAN	17
4.1 <i>Scrapping Data</i>	17
4.2 Pelabelan dan Pemberian <i>Confidence Score</i> pada Data	17
4.3 Pemeriksaan Validitas Label	18
4.4 Eksplorasi Data	18
4.5 <i>Preprocessing Data</i>	20
4.6 Pemodelan dan Evaluasi Model <i>Multinomial Naïve Bayes</i>	21
4.7 Eksplorasi Hasil	23
V SIMPULAN DAN SARAN	30
5.1 Simpulan	30
5.2 Saran	30
DAFTAR PUSTAKA	31
LAMPIRAN	35
RIWAYAT HIDUP	38

DAFTAR TABEL

1	Contoh <i>output</i> dari GPT-4o mini	11
2	Nilai hiperparameter	15
3	Data sebelum dan sesudah tahap <i>preprocessing</i> data	20
4	Performa model klasifikasi	21
5	Performa model MNB dengan <i>back-translation</i>	22
6	<i>Hyperparameter</i> terbaik model MNB dengan <i>back-translation</i>	22
7	Uraian dan contoh kata pada kelas sentimen positif	24
8	Uraian dan contoh kata pada kelas sentimen negatif	25
9	Uraian dan contoh kata pada kelas sentimen netral	26

DAFTAR GAMBAR

1	<i>confussion matrix</i> (a) dan <i>multiclass confussion matrix</i> (b)	8
2	Diagram alir penelitian	10
3	Histogram persebaran <i>confidence score</i> pada data	17
4	Perubahan komposisi sentimen pada keseluruhan dokumen pada bulan Januari (a) & April (b)	19
5	Perubahan komposisi favorit pada semua sentimen pada bulan Januari (a) & April (b)	19
6	<i>Precision-recall curve</i>	23
7	<i>Sentiment wordcloud</i> berdasarkan frasa untuk kelas sentimen positif	24
8	<i>Sentiment wordcloud</i> berdasarkan frasa untuk kelas sentimen negatif	26
9	<i>Sentiment wordcloud</i> berdasarkan frasa untuk kelas sentimen netral	27
10	<i>Sentiment wordcloud</i> berdasarkan frasa untuk kelas sentimen positif (a), netral (b), dan negatif (c) untuk komentar di bulan Januari	28
11	<i>Sentiment Wordcloud</i> berdasarkan frasa untuk kelas sentimen positif(a), netral (b), dan negatif (c) untuk komentar di bulan April	29

DAFTAR LAMPIRAN

1	Uraian performa model MNB dengan metode <i>undersampling</i>	36
2	Uraian performa model MNB dengan tidak menggunakan metode <i>balancing</i>	36
3	Uraian performa model MNB dengan metode <i>back-translation</i>	36
4	Uraian performa model MNB dengan metode SMOTE	36
5	Contoh hasil pelabelan oleh LLM	37