



**Klasifikasi Sentimen dan Pemodelan Topik  
pada Program Makan Bergizi Gratis dengan  
Algoritma Naïve Bayes, KNN, dan BERTopic**

**Disusun oleh:**

**Drinky Ferdianoric**

**41521010193**

**UNIVERSITAS  
MERCU BUANA**

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS MERCU BUANA  
JAKARTA  
2025**



**Klasifikasi Sentimen dan Pemodelan Topik  
pada Program Makan Bergizi Gratis dengan  
Algoritma Naïve Bayes, KNN, dan BERTopic**

**Disusun oleh:**

**Drinky Ferdianoric**

**41521010193**



**Diajukan sebagai salah satu syarat untuk memperoleh gelar sarjana**

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS MERCU BUANA  
JAKARTA  
2025**

## HALAMAN PERNYATAAN KARYA SENDIRI

Saya yang bertanda tangan di bawah ini:

Nama : DRINKY FERDIANORIC  
NIM : 41521010193  
Program Studi : Teknik Informatika  
Judul Proposal Penelitian : Analisa Sentimen Masyarakat Klasifikasi Sentimen dan Pemodelan Topik pada Program Makan Bergizi Gratis dengan Algoritma Naïve Bayes, KNN, dan BERTopic

Menyatakan bahwa Laporan Skripsi ini adalah hasil karya saya sendiri dan bukan plagiat, serta semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar. Apabila ternyata ditemukan di dalam Proposal Penelitian saya terdapat unsur plagiat, maka saya siap mendapatkan sanksi akademis yang berlaku di Universitas Mercu Buana.

Jakarta, 27 Juli 2025



Drinky Ferdianoric

UNIVERSITAS  
MERCU BUANA

## HALAMAN PENGESAHAN

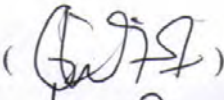
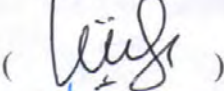
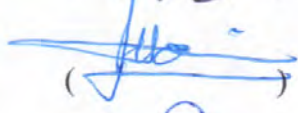
Laporan Skripsi ini diajukan oleh:

Nama : DRINKY FERDIANORIC  
NIM : 41521010193  
Program Studi : Teknik Informatika  
Klasifikasi Sentimen dan Pemodelan Topik pada  
Judul Proposal Penelitian : Program Makan Bergizi Gratis dengan Algoritma  
Naïve Bayes, KNN, dan BERTopic

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 Teknik Informatika, Fakultas Ilmu Komputer Universitas Mercu Buana.

Disahkan oleh:

Pembimbing : Wawan Gunawan, S.Kom., M.T., M.Kom.  
NIDN : 0424108104  
Ketua Penguji : Inna Sabily Karima, S.Kom., M.Kom.  
NIDN : 0324018902  
Penguji 1 : Lukman Hakim, S.T., M.Kom.  
NIDN : 0327107701  
Penguji 2 : Prastika Indriyanti, S.Kom, MCS  
NIDN : 0312089401

()  
()  
()  
()

Jakarta, 4 Agustus 2025

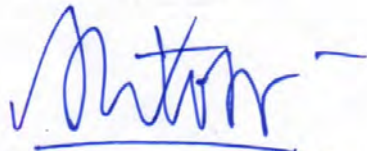
Mengetahui,

Dekan

Ketua Program Studi



Dr. Bambang Jokonowo, S.Si., MTI  
NIDN : 0320037002



Dr. Hadi Santoso, S.Kom., M.Kom  
NIDN : 0225067701

## KATA PENGANTAR

Puji syukur kehadiran Tuhan yang Maha Esa, atas segala rahmat dan ridha-Nya sehingga penulis dapat menyelesaikan proposal penelitian yang merupakan salah satu persyaratan kelulusan Program Studi Strata Satu (S1) pada jurusan Teknik Informatika, Universitas Mercu Buana.

Penulis menyadari bahwa proposal penelitian ini masih jauh dari sempurna, karena kesempurnaan sejatinya hanya milik Tuhan yang Maha Esa. Oleh karena itu, saran dan masukan yang membangun senantiasa penulis terima dengan senang hati. Serta berkat dukungan, motivasi, bantuan, bimbingan, dan doa dari banyak pihak, penulis mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. Andi Adriansyah, M.Eng. selaku Rektor Universitas Mercu Buana.
2. Bapak Dr. Bambang Jokonowo, S.Si., MTI selaku Dekan Fakultas Ilmu Komputer.
3. Bapak Dr. Hadi Santoso, S.Kom., M.Kom. selaku Ketua Program Studi Teknik Informatika Universitas Mercubuanya.
4. Bapak Wawan Gunawan, S.Kom., M.T. selaku dosen pembimbing tugas akhir yang telah memberikan pengarahan, motivasi, menyediakan waktu, tenaga, dan pikiran sehingga selama pembuatan proposal penelitian ini terjadwal dengan baik.
5. Kedua Orang Tua saya yang selalu mensupport dan mendukung saya selama menjalani masa studi sebagai mahasiswa Universitas Mercubuanya..
6. Semua teman kuliah yang selalu berbagi informasi dan memberikan dukungan dalam bentuk yang berbeda-beda.

Akhir kata, penulis berharap semoga Tuhan yang Maha Esa membalas kebaikan dan selalu mencurahkan rahmat, hidayah, serta panjang umur kepada kita semua, aamiin. Terima Kasih.

Jakarta, 2 Juni 2025



Drinky Ferdianoric

## HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

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

Nama : Drinky Ferdianoric  
NIM : 41521010193  
Program Studi : Teknik Informatika  
Analisa Sentimen Masyarakat Klasifikasi  
Judul Proposal Penelitian : Sentimen dan Pemodelan Topik pada Program  
Makan Bergizi Gratis dengan Algoritma Naïve  
Bayes, KNN, dan BERTopic

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 Laporan Magang/Skripsi/Tesis/Disertasi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian Pernyataan ini saya buat dengan sebenarnya.

UNIVERSITAS  
MERCU BUANA

Jakarta, 27 Juli 2025



Drinky Ferdianoric

## ABSTRAK

Nama : Drinky Ferdianoric  
NIM : 41521010193  
Program Studi : Teknik Informatika  
Judul Proposal Penelitian : Program Makan Bergizi Gratis dengan Algoritma Naïve Bayes, KNN, dan BERTopic  
Dosen Pembimbing : Wawan Gunawan, S.Kom., M.T

Penelitian ini menganalisis sentimen masyarakat terhadap Program Makan Bergizi Gratis (MBG), suatu inisiatif gizi nasional yang digagas oleh Presiden Republik Indonesia, Prabowo Subianto, dan dilaksanakan oleh Badan Gizi Nasional (BGN) sebagai koordinator utama. Program ini didukung oleh Kementerian Pendidikan Dasar dan Menengah (Kemendikdasmen) yang menyusun pedoman bagi satuan pendidikan, Kementerian Kesehatan (Kemenkes) yang bertanggung jawab pada aspek kesehatan dan gizi, serta Bappenas dan UNICEF yang berperan dalam penyusunan rekomendasi tata kelola dan sistem evaluasi.

Data opini publik dikumpulkan dari media sosial Twitter, kemudian melalui proses pembersihan teks, penerjemahan, dan analisis topik menggunakan BERTopic untuk mengidentifikasi isu-isu dominan. Analisis sentimen dilakukan dengan algoritma Naïve Bayes dan K-Nearest Neighbor (KNN) untuk mengklasifikasikan opini menjadi sentimen positif, negatif, dan netral. Tujuan penelitian ini adalah mengukur persepsi publik secara otomatis dan sistematis, serta membandingkan kinerja kedua algoritma. Hasil penelitian diharapkan dapat menjadi bahan pertimbangan bagi pengambil kebijakan dalam evaluasi program dan menjadi referensi bagi akademisi maupun praktisi dalam penerapan machine learning dan topic modeling pada kebijakan publik.

**Kata kunci:** Analisis Sentimen, Makan Bergizi Gratis, Prabowo Subianto, Naïve Bayes, K-Nearest Neighbor, BERTopic, Twitter

## ABSTRACT

Nama : Drinky Ferdianoric  
NIM : 41521010193  
Program Studi : Teknik Informatika  
Klasifikasi Sentimen dan Pemodelan Topik pada  
Judul Proposal Penelitian : Program Makan Bergizi Gratis dengan Algoritma  
Naïve Bayes, KNN, dan BERTopic  
Dosen Pembimbing : Wawan Gunawan, S.Kom., M.T

This study analyzes public sentiment toward the Free Nutritious Meal Program (Makan Bergizi Gratis/MBG), a national nutrition initiative initiated by the President of the Republic of Indonesia, Prabowo Subianto, and coordinated by the National Nutrition Agency (Badan Gizi Nasional/BGN) as the main implementing body. The program is supported by the Ministry of Primary and Secondary Education (Kementerian Pendidikan Dasar dan Menengah/Kemendikdasmen), which develops implementation guidelines for educational institutions; the Ministry of Health (Kementerian Kesehatan/Kemenkes), which is responsible for health and nutritional aspects; as well as Bappenas and UNICEF, which contribute to governance recommendations and evaluation systems.

Public opinion data were collected from the social media platform Twitter, then processed through text cleaning, translation, and topic analysis using the BERTopic method to identify dominant issues. Sentiment analysis was conducted using Naïve Bayes and K-Nearest Neighbor (KNN) algorithms to classify opinions into positive, negative, and neutral categories. The aim of this research is to measure public perception in an automated and systematic manner, as well as to compare the performance of the two algorithms. The findings are expected to provide valuable input for policymakers in evaluating the program and to serve as a reference for academics and practitioners in applying machine learning and topic modeling to public policy analysis.

**Keyword:** Analisis Sentimen, Makan Bergizi Gratis, Prabowo Subianto, Naïve Bayes, K-Nearest Neighbor, BERTopic, Twitter

## DAFTAR ISI

|                                                                                                 |             |
|-------------------------------------------------------------------------------------------------|-------------|
| <b>HALAMAN JUDUL .....</b>                                                                      | <b>i</b>    |
| <b>HALAMAN PERNYATAAN KARYA SENDIRI .....</b>                                                   | <b>ii</b>   |
| <b>HALAMAN PENGESAHAN.....</b>                                                                  | <b>iii</b>  |
| <b>KATA PENGANTAR.....</b>                                                                      | <b>iv</b>   |
| <b>HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS<br/>AKHIR UNTUK KEPENTINGAN AKADEMIS.....</b> | <b>v</b>    |
| <b>ABSTRAK .....</b>                                                                            | <b>vi</b>   |
| <b>ABSTRACT.....</b>                                                                            | <b>vii</b>  |
| <b>DAFTAR ISI .....</b>                                                                         | <b>viii</b> |
| <b>DAFTAR TABEL .....</b>                                                                       | <b>xi</b>   |
| <b>DAFTAR GAMBAR.....</b>                                                                       | <b>xii</b>  |
| <b>DAFTAR LAMPIRAN.....</b>                                                                     | <b>xii</b>  |
| <b>BAB I PENDAHULUAN.....</b>                                                                   | <b>xiii</b> |
| 1.1 Latar Belakang.....                                                                         | 1           |
| 1.2 Identifikasi Masalah.....                                                                   | 3           |
| 1.3 Batasan Masalah .....                                                                       | 3           |
| 1.4 Tujuan Penelitian .....                                                                     | 4           |
| 1.5 Manfaat Penelitian .....                                                                    | 4           |
| <b>BAB II TINJAUAN PUSTAKA .....</b>                                                            | <b>6</b>    |
| 2.1 Twitter.....                                                                                | 6           |
| 2.2 Crawling.....                                                                               | 6           |
| 2.3 Text Mining .....                                                                           | 7           |
| 2.4 Algoritma .....                                                                             | 7           |
| 2.4.1 Naive Bayes .....                                                                         | 7           |
| 2.4.2 K Nearest Neighbor .....                                                                  | 8           |
| 2.5 Sentiment Analysis .....                                                                    | 9           |
| 2.6 Pemodelan Topik .....                                                                       | 10          |
| 2.7 Literatur Review .....                                                                      | 10          |
| <b>BAB III METODE RISET .....</b>                                                               | <b>16</b>   |
| 3.1 Metodologi Penelitian.....                                                                  | 16          |

|                                                       |           |
|-------------------------------------------------------|-----------|
| 3.2 Subjek dan Objek Penelitian .....                 | 18        |
| 3.3 Tahapan Penelitian.....                           | 18        |
| 3.3.1 Studi literatur .....                           | 18        |
| 3.3.2 Pengumpulan Data.....                           | 22        |
| 3.3.3 Pre-processing data .....                       | 23        |
| 3.3.4 Labelling .....                                 | 25        |
| 3.3.5 Data Training dan Data Testing.....             | 26        |
| 3.3.6 Visualization dan Data Comparison .....         | 27        |
| 3.3.7 Pemodelan Topik BERTopic .....                  | 28        |
| <b>BAB IV PEMBAHASAN .....</b>                        | <b>30</b> |
| 4.1 Preprocessing .....                               | 30        |
| 4.1.1 Cleaning .....                                  | 30        |
| 4.1.2 Case Folding .....                              | 33        |
| 4.1.3 Normalize text.....                             | 36        |
| 4.1.4 Stemmed .....                                   | 38        |
| 4.1.5 Tokenization .....                              | 40        |
| 4.1.6 Stop Word Removal.....                          | 42        |
| 4.2 Labeling TextBlob .....                           | 44        |
| 4.3 Analysis .....                                    | 48        |
| 4.3.1 Konversi Label Sentimen (Convert Labeling)..... | 48        |
| 4.3.2 Splitting Data .....                            | 51        |
| 4.4 Visualisasi .....                                 | 57        |
| 4.4.1 WordCloud.....                                  | 58        |
| 4.4.2 Barplot .....                                   | 60        |
| 4.5 Pemodelan Topik .....                             | 62        |
| 4.5.1 Hasil dari BERTopic modeling.....               | 68        |
| <b>BAB V KESIMPULAN DAN SARAN.....</b>                | <b>78</b> |
| 5.1 Kesimpulan .....                                  | 78        |
| 5.2 Saran .....                                       | 79        |

|                                       |           |
|---------------------------------------|-----------|
| 4.5.2 Hasil dari Survey Lapangan..... | 79        |
| <b>DAFTAR PUSTAKA.....</b>            | <b>81</b> |
| <b>LAMPIRAN .....</b>                 | <b>86</b> |



## DAFTAR TABEL

|                                                       |    |
|-------------------------------------------------------|----|
| Tabel 1 Literature Review .....                       | 11 |
| Tabel 2 Spesifikasi .....                             | 29 |
| Tabel 3 Hasil Cleaning .....                          | 32 |
| Tabel 4 Hasil Case Folding.....                       | 35 |
| Tabel 5 Hasil Normalize Text.....                     | 37 |
| Tabel 6 Hasil Stemmed.....                            | 39 |
| Tabel 7 Hasil Tokenization.....                       | 41 |
| Tabel 8 Hasil Stop Word Removal .....                 | 43 |
| Tabel 9 Hasil Labeling TextBlob.....                  | 48 |
| Tabel 10 Hasil Convert Labeling.....                  | 54 |
| Tabel 11 Hasil Jumlah Data Hasil Pre Processing ..... | 55 |
| Tabel 12 Splitting Naïve Bayes 60:40 .....            | 55 |
| Tabel 13 Splitting K-Nearest Neighbor 60:40 .....     | 56 |
| Tabel 14 Splitting Naïve Bayes 70:30 .....            | 57 |
| Tabel 15 Splitting K-Nearest Neighbor 70:30 .....     | 58 |
| Tabel 16 Splitting Naïve Bayes 80:20 .....            | 59 |
| Tabel 17 Splitting K-Nearest Neighbor 80:20 .....     | 59 |
| Tabel 18 Hasil BERTopic.....                          | 68 |



## DAFTAR GAMBAR

|                                                |    |
|------------------------------------------------|----|
| Gambar 3.1.1 Tahapan penelitian .....          | 18 |
| Gambar 4.4.1.1 WordCloud Positif .....         | 63 |
| Gambar 4.4.1.2 WordCloud Negatif.....          | 68 |
| Gambar 4.4.2.1 Diagram Barplot Tweet MBG.....  | 66 |
| Gambar 4.5.2.1 Diagram Persentase respon ..... | 81 |



## DAFTAR LAMPIRAN

|                                           |    |
|-------------------------------------------|----|
| Lampiran 1 Kartu Asistensi .....          | 86 |
| Lampiran 2 Curriculum Vitae .....         | 87 |
| Lampiran 3 Surat Pernyataan HAKI.....     | 89 |
| Lampiran 4 Sertifikat BNSP .....          | 92 |
| Lampiran 5 Form Revisi Dosen Penguji..... | 93 |
| Lampiran 6 Hasil Cek Plagiat .....        | 95 |
| Lampiran 7 Halaman Persetujuan.....       | 96 |

