

**ANALISIS SENTIMEN OPINI PUBLIK TERHADAP
PROGRAM MAKAN BERGIZI GRATIS DI MEDIA SOSIAL X
MENGUNAKAN NAÏVE BAYES CLASSIFIER (NBC) DAN
SUPPORT VECTOR MACHINE (SVM)**

LAPORAN TUGAS AKHIR



**LAISA USRAINI
1222001021**

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN ILMU KOMPUTER
UNIVERSITAS BAKRIE
JAKARTA
2026**

HALAMAN PERNYATAAN ORISINALITAS

Tugas Akhir ini adalah hasil karya saya sendiri, dan semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar.

Nama : Laisa Ussaini

NIM : 1222001021

Tanda Tangan : 

Tanggal : 15 September 2026

HALAMAN PENGESAHAN

Tugas Akhir ini diajukan oleh :

Nama : Laisa Usraini
NIM : 1222001021
Program Studi : Teknik Informatika
Fakultas : Fakultas Teknik dan Ilmu Komputer
Judul Penelitian : Analisis Sentimen Opini Publik terhadap Program Makan Bergizi Gratis di Media Sosial X Menggunakan *Naive Bayes Classifier* (NBC) dan *Support Vector Machine* (SVM)

Telah berhasil dipertahankan di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Komputer pada Program Studi Informatika, Fakultas Teknik dan Ilmu Komputer, Universitas Bakrie.

DEWAN PENGUJI

Penguji 1: Dimas Aryo Anggoro, S.Kom., M.Sc.

18/09/2026 9:10 WIB

(.....)

Penguji 2: Dewi Fatmawati Surianto, S.Kom., M.Kom.

18/09/2026 11:25 WIB

(.....)

Pembimbing: Ir. Iwan Adhicandra, S.T., M.Sc., Ph.D.,
SMIEEE, MACM, MIET, MBCS

(.....)

Ditetapkan di : Jakarta

Tanggal : 15 September 2026

KATA PENGANTAR

Puji syukur penulis panjatkan kepada Tuhan Yang Maha Esa, karena atas berkat dan rahmat-Nya, penulis dapat menyusun Tugas Akhir ini yang berjudul “Analisis Sentimen Opini Publik terhadap Program Makan Bergizi Gratis di Media Sosial X Menggunakan *Naive Bayes Classifier* (NBC) dan *Support Vector Machine* (SVM)” sebagai salah satu syarat untuk mencapai gelar Sarjana Komputer Program Studi Informatika, Fakultas Teknik dan Ilmu Komputer, Universitas Bakrie. Dalam proses penyusunan tugas akhir ini, penulis banyak mendapatkan bantuan, bimbingan, serta dukungan dari berbagai pihak. Oleh karena itu, penulis mengucapkan terima kasih:

- 1) Kepada pemilik mutlak seluruh zat dan alam semesta, yang menguasai, mengatur, dan berkuasa atas segala sesuatu. Allah SWT, Sang Maha Pemilik Arah. Terima kasih telah menjadi satu-satunya rumah tempatku pulang saat dunia terasa begitu kabur dan bising. Terima kasih telah menjaga waras dan langkahku, bahkan disaat penulis hampir lupa alasan untuk hidup.
- 2) Bapak Ir. Iwan Adhicandra, S.T., M.Sc., Ph.D, selaku Ketua Program Studi dan sekaligus Dosen Pembimbing penulis yang telah bersedia meluangkan waktu, memberikan arahan, bimbingan, masukan bernilai, serta kesabaran yang luar biasa dalam mendampingi penulis hingga tugas akhir ini selesai.
- 3) Seluruh dosen Universitas Bakrie, khususnya di Program Studi Informatika, yang telah memberikan bekal ilmu pengetahuan, bimbingan, dan wawasan selama penulis menempuh masa perkuliahan.
- 4) Kedua orang tua tercinta, Papa Mukhtar dan Mama Rosmita, yang merupakan alasan terbesar dibalik setiap langkah dan pencapaian penulis hingga saat ini. Terima kasih atas cinta, kasih sayang, dan doa yang tidak pernah putus. Gelar sarjana ini bukanlah akhir perjuangan, melainkan salah satu bentuk bakti dan rasa terima kasih penulis atas setiap air mata, doa, kerja keras, dan cinta yang telah mengantarkan penulis sampai pada titik ini. Semoga pencapaian ini menjadi awal untuk membahagiakan serta membalas, meskipun tidak akan pernah sebanding dengan seluruh pengorbanan yang telah beliau berikan.

- 5) Kepada Kakak tercinta, Kakak Liza yang tidak hanya menjadi seorang kakak, tetapi juga panutan, tempat berbagi cerita serta sosok yang memiliki peran besar dalam membentuk penulis menjadi pribadi seperti ini. Terima kasih selalu mengulurkan tangan, memberikan hiburan, dorongan semangat, dan bantuan kecil yang sangat berarti di sela-sela proses pengerjaan tugas akhir ini. Kepada Adik tercinta Najwa, dibalik segala kekerakepalaan dan tingkah laku yang menguji kesabaran, ia adalah sosok yang mengajari penulis banyak hal tentang arti menerima dan menyayangi.
- 6) Terima kasih kepada seluruh keluarga besar atas doa dan dukungan yang telah diberikan selama proses perkuliahan.
- 7) Kepada Rosvita, saudara sekaligus sahabat yang selalu ada ketika penulis membutuhkan telinga untuk mendengarkan semua keluhan kesah. terima kasih atas candaan, dukungan, dan semangat yang selalu diberikan kepada penulis.
- 8) Aisyifa, sahabat penulis yang telah menjadi teman seperjuangan penulis sejak sekolah menengah pertama hingga hari ini yang selalu memberikan semangat dan doa baik kepada penulis.
- 9) Teman-teman seperjuangan Teknik Informatika, terima kasih atas tawa, semangat, dan kebersamaan selama masa perkuliahan
- 10) Seluruh pihak yang telah memberikan dukungan, bantuan, dan doa selama proses penyusunan tugas akhir ini, yang tidak dapat penulis sebutkan satu per satu.

Semoga Allah SWT membalas segala kebaikan dunia dan akhirat kepada semua pihak yang telah membantu. Penulis menyadari bahwa penulisan Tugas Akhir ini masih jauh dari kata sempurna, penulis berharap bahwa penelitian dan Tugas Akhir ini dapat memberikan manfaat dan berguna bagi kalangan Pendidikan. Khususnya bidang Informatika.

Jakarta, 15 September 2026

A handwritten signature in black ink, appearing to read 'Laisa Ussaini'.

Laisa Ussaini

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI

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

Nama : Laisa Usraini
NIM : 1222001021
Program Studi : Informatika
Fakultas : Fakultas Teknik dan Ilmu Komputer

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Bakrie **Hak Bebas Royalti Noneksklusif** (*Non-exclusive Royalty-Free Right*) atas karya ilmiah saya yang berjudul:

Analisis Sentimen Opini Publik terhadap Program Makan Bergizi Gratis di Media Sosial X Menggunakan *Naive Bayes Classifier* (NBC) dan *Support Vector Machine* (SVM)

beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Noneksklusif ini Universitas Bakrie berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta untuk kepentingan akademis.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta
Pada tanggal : 15 September 2026

Yang menyatakan



(Laisa Usraini)

Analisis Sentimen Opini Publik terhadap Program Makan Bergizi Gratis di Media Sosial X Menggunakan Naïve Bayes Classifier (NBC) dan Support Vector Machine (SVM)

Laisa Usraini

ABSTRAK

Program Makan Bergizi Gratis (MBG) merupakan salah satu program prioritas pemerintah yang memicu banyak tanggapan masyarakat di media sosial X. Besarnya jumlah opini membuat analisis manual menjadi kurang efektif, sekaligus penelitian terdahulu umumnya hanya membandingkan akurasi tanpa mengukur efisiensi komputasi secara kuantitatif. Penelitian ini bertujuan membandingkan kinerja klasifikasi sekaligus efisiensi komputasi algoritma *Naïve Bayes Classifier* (NBC) dan *Support Vector Machine* (SVM) dalam menganalisis sentimen opini publik terhadap program MBG. Metode yang digunakan mengikuti kerangka kerja SEMMA, mencakup pengumpulan tweet berbahasa Indonesia melalui Tweet-Harvest, *preprocessing*, pelabelan sentimen, dan pembobotan fitur dengan TF-IDF. Kedua algoritma diuji pada rasio 70:30, 80:20, dan 90:10 serta divalidasi dengan *K-Fold Cross Validation* ($k=10$), lalu dievaluasi berdasarkan *accuracy*, *precision*, *recall*, dan *F1-score*, serta dibandingkan dari sisi training time dan inference time. Hasil penelitian menunjukkan SVM memperoleh kinerja klasifikasi lebih baik dibandingkan NBC, dengan rata-rata akurasi 73,18% dan F1-makro 67,35% pada validasi K-Fold Cross Validation ($K=10$), sedangkan NBC memperoleh akurasi 62,74% dan F1-makro 47,16%. Sebaliknya, NBC jauh lebih unggul dari sisi efisiensi, karena SVM membutuhkan waktu pelatihan dan prediksi hingga ratusan kali lebih lama. Penelitian ini menyimpulkan terdapat trade-off antara kinerja klasifikasi dan efisiensi komputasi, sehingga pemilihan algoritma perlu mempertimbangkan keduanya sesuai kebutuhan.

Kata kunci: analisis sentimen, Makan Bergizi Gratis, Naïve Bayes Classifier, Support Vector Machine, efisiensi komputasi

Public Sentiment Analysis on the Free Nutritious Meal Program on Social Media X Using Naïve Bayes Classifier (NBC) and Support Vector Machine (SVM)

Laisa Usraini

ABSTRACT

The Free Nutritious Meal (MBG) program is one of the government's priority initiatives that has generated significant public response on the social media platform X. The massive volume of opinions renders manual analysis ineffective, while previous studies generally focused only on comparing accuracy without quantitatively measuring computational efficiency. This study aims to evaluate and compare the classification performance alongside the computational efficiency of the Naïve Bayes Classifier (NBC) and Support Vector Machine (SVM) algorithms in analyzing public sentiment toward the MBG program. The methodology follows the SEMMA framework, encompassing the collection of Indonesian-language tweets via Tweet-Harvest, preprocessing, sentiment labeling, and feature weighting using TF-IDF. Both algorithms were tested across data split ratios of 70:30, 80:20, and 90:10, and validated using 10-Fold Cross Validation ($k=10$). Evaluation metrics included accuracy, precision, recall, and F1-score, as well as comparative training time and inference time. The results indicate that SVM achieved superior classification performance compared to NBC, attaining an average accuracy of 73.18% and a macro F1-score of 67.35% under 10-Fold Cross Validation, whereas NBC recorded an accuracy of 62.74% and a macro F1-score of 47.16%. Conversely, NBC demonstrated significantly higher efficiency, as SVM required hundreds of times longer for both training and inference. This study concludes that a clear trade-off exists between classification performance and computational efficiency, highlighting the need to select algorithms based on specific operational requirements and constraints.

Keywords: sentiment analysis, Free Nutritious Meals, Naïve Bayes Classifier, Support Vector Machine, computational efficiency

DAFTAR ISI

HALAMAN PERNYATAAN ORISINALITAS	i
HALAMAN PENGESAHAN.....	ii
KATA PENGANTAR.....	iii
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI	vi
ABSTRAK	vii
ABSTRACT.....	viii
DAFTAR ISI	ix
DAFTAR GAMBAR	xii
DAFTAR TABEL.....	xiv
DAFTAR SINGKATAN	xvi
BAB 1. PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan Penelitian.....	4
1.4 Manfaat Penelitian	4
1.5 Batasan Penelitian	5
1.6 Sistematika Penulisan.....	5
BAB 2. TINJAUAN PUSTAKA.....	7
2.1 Landasan Teori	7
2.1.1 Analisis Sentimen.....	7
2.1.2 <i>Data Mining</i>	8
2.1.3 Media Sosial X (Twitter).....	9
2.1.4 Makan Bergizi Gratis	10
2.1.5 <i>Preprocessing Data</i>	11
2.1.6 Pelabelan Data.....	12
2.1.7 Ketidakseimbangan Data, Random Sampling, dan Penentuan Ukuran Sample.....	15
2.1.8 Algoritma <i>Naïve Bayes Classifier</i>	16
2.1.9 Algoritma <i>Support Vector Machine</i> (SVM)	17
2.1.10 Metode SEMMA.....	20
2.1.11 <i>Confusion Matrix</i>	23
2.1.12 <i>Term Frequency-Inverse Document Frequency</i> (TF-IDF)	25

2.1.13	<i>K-Fold Cross Validation</i>	27
2.1.14	<i>Word Cloud</i>	28
2.1.15	Efisiensi Komputasi	29
2.1.16	Kompleksitas Algoritma.....	30
2.1.17	Python	32
2.1.18	<i>Google Colaboratory</i>	33
2.2	Penelitian Terdahulu.....	34
BAB 3. METODE PENELITIAN.....		40
3.1	Jenis Penelitian.....	40
3.2	Sumber Data.....	40
3.3	Tahapan Penelitian	42
3.4	<i>Sample</i>	44
3.5	<i>Explore</i>	44
3.6	<i>Modify</i>	45
3.6.1	<i>Cleansing</i>	45
3.6.2	<i>Case Folding</i>	46
3.6.3	Tokenisasi.....	47
3.6.4	Normalisasi	48
3.6.5	<i>Stopword Removal</i>	49
3.6.6	<i>Stemming</i>	50
3.7	Model	53
3.7.1	Pelabelan Data.....	53
3.7.2	Pembobotan Fitur TF-IDF.....	57
3.7.3	Pembagian Data	58
3.7.4	Klasifikasi Model.....	60
3.7.5	K-Fold Cross Validation.....	61
3.7.6	Pengukuran Efisiensi Komputasi	62
3.8	Assess.....	63
3.8.1	Evaluasi Performa Klasifikasi.....	64
3.8.2	Perbandingan Efisiensi Komputasi	65
3.8.3	Visualisasi Word Cloud.....	66
3.9	Perangkat Penelitian.....	66

BAB 4 HASIL DAN PEMBAHASAN.....	68
4.1 Sample.....	68
4.1.1 Pengumpulan Data (<i>Scraping</i>).....	68
4.2 Explore.....	69
4.3 Modify.....	70
4.3.1 Cleansing.....	71
4.3.2 <i>Case Folding</i>	73
4.3.3 Tokenisasi.....	74
4.3.4 Normalisasi	75
4.3.5 <i>Stopword Removal</i>	77
4.3.6 Stemming	80
4.4 Model	82
4.4.1 Pelabelan Data.....	82
4.4.2 Pembagian Data	89
4.4.3 Pembobotan TF-IDF	92
4.4.4 Klasifikasi <i>Naïve Byes Classifier</i>	96
4.4.5 Klasifikasi <i>Support Vector Machine</i>	100
4.4.6 Validasi <i>K-Fold Cross Validation</i>	104
4.5 Assess	108
4.5.1 Perbandingan Kinerja NBC dan SVM	108
4.5.2 Perbandingan Efisiensi Komputasi NBC dan SVM.....	115
4.5.3 Uji Gold Standard: Perbandingan Model dengan Penilaian Manual ..	120
4.6 Pembahasan.....	126
4.6.1 Interpretasi Hasil	127
4.6.2 Keterkaitan dengan Kompleksitas Algoritma	129
4.6.3 Keterbatasan Penelitian.....	130
BAB 5 KESIMPULAN DAN SARAN.....	133
5.1 Kesimpulan	133
5.2 Saran.....	134
DAFTAR PUSTAKA	136
LAMPIRAN	146

DAFTAR GAMBAR

Gambar 2. 1 skematis proses klasifikasi hyperplane pada algoritma SVM.....	18
Gambar 2. 2 Tahapan Metode SEMMA.....	21
Gambar 2. 3 Skema K-Fold Cross Validation	28
Gambar 3. 1 Alur Tahapan Penelitian	43
Gambar 3. 2 Flowchart Cleansing.....	46
Gambar 3. 3 Flowchart Case Folding	47
Gambar 3. 4 Flowchart Tokenisasi.....	48
Gambar 3. 5 Flowchart Normalisasi	49
Gambar 3. 6 Flowchart Stopword Removal.....	50
Gambar 3. 7 Flowchart Stemming	51
Gambar 4. 1 Proses Scraping	68
Gambar 4. 2 Hasil penyaringan atribut	70
Gambar 4. 3 Proses Cleansing	72
Gambar 4. 4 Proses Case Folding	74
Gambar 4. 5 Proses Tokenisasi	75
Gambar 4. 6 Proses Normalisasi	77
Gambar 4. 7 Proses Stopword Removal	79
Gambar 4. 8 Proses Stemming.....	81
Gambar 4. 9 Distribusi label sentimen.....	89
Gambar 4. 10 Perbandingan Performa NBC pada Kondisi Imbalanced dan Balanced.....	98
Gambar 4. 11 Waktu Pelatihan dan Prediksi NBC pada Kondisi Imbalanced dan Balanced.....	100
Gambar 4. 12 Perbandingan Performa SVM pada Kondisi Imbalanced dan Balanced.....	101
Gambar 4. 13 Waktu Pelatihan dan Prediksi SVM pada Kondisi Imbalanced dan Balanced.....	103
Gambar 4. 14 Perubahan Rata-Rata Akurasi NBC dan SVM terhadap Nilai K .	105
Gambar 4. 15 Perbandingan Kinerja NBC dan SVM pada Kondisi Imbalanced dan Balanced.....	109
Gambar 4. 16 Perbandingan Rata-Rata Akurasi dan F1-Makro pada K = 10.....	111
Gambar 4. 17 Confusion Matrix NBC pada Rasio 90:10 Kondisi Imbalanced...	113
Gambar 4. 18 Confusion Matrix SVM pada Rasio 90:10 Kondisi Imbalanced...	114
Gambar 4. 19 Perbandingan Waktu Pelatihan dan Prediksi NBC dan SVM pada Skala Logaritmik.....	116
Gambar 4. 20 Perbandingan Waktu Total Cross-Validation NBC dan SVM pada Skala Logaritmik.....	118
Gambar 4. 21 Perbandingan Akurasi InSet Refined, NBC, dan SVM terhadap Label Manual	121
Gambar 4. 22 Word cloud entimen positif.....	123
Gambar 4. 23 Word cloud sentimen negatif.....	124

Gambar 4. 24 Word cloud sentimen netral.....	125
--	-----

DAFTAR TABEL

Tabel 2. 1 Langkah-langkah kesepakatan untuk data kategorikal.....	14
Tabel 2. 2 Confusion Matrix	23
Tabel 2. 3 Perbandingan Kompleksitas Algoritma NBC dan SVM	31
Tabel 2. 4 Penelitian Terdahulu.....	37
Tabel 3. 1 Contoh Transformasi Data pada Setiap Tahap Preprocessing	52
Tabel 3. 2 Contoh Keterbatasan Normalisasi pada Ekspresi Tawa	53
Tabel 3. 3 Perbandingan Distribusi Kelas pada Beberapa Nilai Ambang.....	55
Tabel 3. 4 Spesifikasi Hardware dan Software	66
Tabel 4. 1 Hasil Scraping tweet.....	69
Tabel 4. 2 Hasil Cleansing	71
Tabel 4. 3 Hasil Case Folding	73
Tabel 4. 4 Hasil Tokenisasi	74
Tabel 4. 5 Hasil Normalisasi	76
Tabel 4. 6 Hasil Stopword Removal	78
Tabel 4. 7 Hasil Stemming	80
Tabel 4. 8 Perubahan Jumlah Data pada Tiap Tahap Preprocessing	82
Tabel 4. 9 Kontribusi Beberapa Kata terhadap Skor Sebelum Pengecualian	83
Tabel 4. 10 Contoh perhitungan skor sentimen positif	84
Tabel 4. 11 Contoh Perhitungan Skor Pelabelan.....	84
Tabel 4. 12 Perbandingan Distribusi Label Sebelum dan Sesudah Refinement ...	85
Tabel 4. 13 Hasil Validasi Pelabelan terhadap Penilaian Manual	86
Tabel 4. 14 Matriks Silang Label Manual dan Label Refined	86
Tabel 4. 15 Distribusi Label Refined pada Dataset Akhir.....	88
Tabel 4. 16 Jumlah Data Latih dan Data Uji pada Setiap Rasio	90
Tabel 4. 17 Distribusi Data Latih Sebelum dan Sesudah Random Undersampling	91
Tabel 4. 18 Contoh Menghitung Term Frequency	92
Tabel 4. 19 Contoh Menghitung Inverse Document Frequency (IDF)	93
Tabel 4. 20 Contoh Menghitung TF-IDF	94
Tabel 4. 21 Jumlah Fitur TF-IDF pada Setiap Skenario Utama.....	96
Tabel 4. 22 Hasil Klasifikasi NBC pada Tiga Rasio Pembagian Data.....	97
Tabel 4. 23 Waktu Komputasi NBC pada Tiga Rasio Pembagian Data.....	99
Tabel 4. 24 Klasifikasi SVM pada Tiga Rasio Pembagian Data.....	101
Tabel 4. 25 Waktu Komputasi SVM pada Tiga Rasio Pembagian Data	103
Tabel 4. 26 Rata-Rata Akurasi dan F1-Makro pada Variasi Nilai K.....	105
Tabel 4. 27 Waktu Total Cross-Validation pada Variasi Nilai K	107
Tabel 4. 28 Perbandingan Kinerja NBC dan SVM pada Tiga Rasio Pembagian Data	109
Tabel 4. 29 Perbandingan Rata-Rata Hasil 10-Fold Cross Validation	111
Tabel 4. 30 Confusion Matrix NBC pada Rasio 90:10 Kondisi Imbalanced	113
Tabel 4. 31 Metrik Tiap Kelas NBC pada Rasio 90:10 Kondisi Imbalanced	114

Tabel 4. 32 Perbandingan Waktu Komputasi NBC dan SVM pada Tiga Rasio...	116
Tabel 4. 33 Perbandingan Waktu Total Cross-Validation NBC dan SVM.....	118
Tabel 4. 34 Hasil Uji Gold Standard	120
Tabel 4. 35 Top 10 kata teratas sentimen positif	123
Tabel 4. 36 Top 10 kata teratas sentimen negatif	124
Tabel 4. 37 Top 10 kata teratas sentimen netral	125

DAFTAR SINGKATAN

API (*Application Programming Interface*)
BERT (*Application Programming Interface*)
BGN (Badan Gizi Nasional)
CPU (*Central Processing Unit*)
CSV (*Comma Separated Values*)
FN (*False Negative*)
FP (*False Positive*)
GPU (*Graphics Processing Unit*)
IDF (*Inverse Document Frequency*)
InSet (*Indonesia Sentiment Lexicon*)
MBG (Makan Bergizi Gratis)
NBC (*Naïve Bayes Classifier*)
NLTK (*Natural Language Toolkit*)
NLP (*Natural Language Processing*)
RAM (*Random Access Memory*)
RBF (*Radial Basis Function*)
SAS (*Statistical Analysis System*)
SEMMA (*Sample, Explore, Modify, Model, Assess*)
SMOTE (*Synthetic Minority Oversampling Technique*)
SVM (*Support Vector Machine*)
TF (*Term Frequency*)
TF-IDF (*Term Frequency-Inverse Document Frequency*)
TN (*True Negative*)
TP (*True Positive*)
TPU (*Tensor Processing Unit*)
URL (*Uniform Resource Locator*)