

**ANALISIS SENTIMEN PROGRAM MAKAN BERGIZI
GRATIS DI MEDIA SOSIAL X DENGAN
STACKING NAIVE BAYES
SVM DAN RF**



SKRIPSI

**AGIL WIRA PRATAMA
00000061000**

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2025**

**ANALISIS SENTIMEN PROGRAM MAKAN BERGIZI
GRATIS DI MEDIA SOSIAL X DENGAN
STACKING NAIVE BAYES
SVM DAN RF**



SKRIPSI

Diajukan sebagai salah satu syarat untuk memperoleh
Gelar Sarjana Komputer (S.Kom.)

**AGIL WIRA PRATAMA
00000061000**

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2025**

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Agil Wira Pratama

Nomor Induk Mahasiswa : 00000061000

Program Studi : Informatika

Skripsi dengan judul:

ANALISIS SENTIMEN PROGRAM MAKAN BERGIZI GRATIS DI MEDIA SOSIAL X DENGAN STACKING NAIVE BAYES SVM DAN RF

merupakan hasil karya saya sendiri bukan plagiat dari laporan karya tulis ilmiah yang ditulis oleh orang lain, dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar serta dicantumkan di Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/penyimpangan, baik dalam pelaksanaan maupun dalam penulisan laporan karya tulis ilmiah, saya bersedia menerima konsekuensi dinyatakan TIDAK LULUS untuk mata kuliah yang telah saya tempuh.

UMM

Tangerang, 1 Juli 2025



(Agil Wira Pratama)

UNIVERSITAS
MULTIMEDIA
NUSANTARA

HALAMAN PENGESAHAN

Skripsi dengan judul

**ANALISIS SENTIMEN PROGRAM MAKAN BERGIZI GRATIS DI MEDIA
SOSIAL X DENGAN STACKING NAIVE BAYES, SVM, RF**

Oleh

Nama : Agil Wira Pratama
NIM : 00000061000
Program Studi : Informatika
Fakultas : Teknik dan Informatika

Telah diujikan pada hari Rabu, 09 Juli 2025

Pukul 13.00 s.d 15.00 dan dinyatakan

LULUS

Dengan susunan penguji sebagai berikut.

Ketua Sidang



Moeljono Widjaja, B.Sc., M.Sc Ph.D.
NIDN: 0311106903

Penguji



Marlinda Vasty Overbeek, S.Kom, M.Kom
NIDN: 0818038501

Pembimbing



David Agustriawan, S.Kom., M.Sc., Ph.D.
NIDN: 0525088601

Ketua Program Studi Informatika



Assoc. Prof. Arya Wicaksana, S.Kom., M.Eng.Sc., OCA
NIDN: 0315109103

**HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK
KEPENTINGAN AKADEMIS**

Yang bertanda tangan di bawah ini:

Nama : Agil Wira Pratama
NIM : 00000061000
Program Studi : Informatika
Jenjang : S1
Judul Karya Ilmiah : ANALISIS SENTIMEN PROGRAM
MAKAN BERGIZI GRATIS DI
MEDIA SOSIAL X DENGAN
STACKING NAIVE BAYES SVM
DAN RF

Menyatakan dengan sesungguhnya bahwa saya bersedia :

- ☒ Saya bersedia memberikan izin sepenuhnya kepada Universitas Multimedia Nusantara untuk mempublikasikan hasil karya ilmiah saya ke dalam repositori Knowledge Center sehingga dapat diakses oleh Sivitas Akademika UMN/Publik. Saya menyatakan bahwa karya ilmiah yang saya buat tidak mengandung data yang bersifat konfidensial.
- ☐ Saya tidak bersedia mempublikasikan hasil karya ilmiah ini ke dalam repositori Knowledge Center, dikarenakan: dalam proses pengajuan publikasi ke jurnal/konferensi nasional/internasional (dibuktikan dengan *letter of acceptance*) **.
- ☐ Lainnya, pilih salah satu:
- Hanya dapat diakses secara internal Universitas Multimedia Nusantara
 - Embargo publikasi karya ilmiah dalam kurun waktu tiga tahun.

Tangerang, 1 Juli 2025

Yang menyatakan



Agil Wira Pratama

HALAMAN PERSEMBAHAN / MOTTO

"Fortis Fortuna Adiuvat"

John Wick



KATA PENGANTAR

Puji syukur atas berkat dan rahmat kepada Tuhan Yang Maha Esa, atas selesainya penulisan laporan tugas akhir dengan judul: Analisis Sentimen Pengguna Media Sosial X terhadap Program Makan Bergizi Gratis Menggunakan Ensemble Learning Stacking dengan Algoritma Naive Bayes, SVM, dan Random Forest, yang dilakukan untuk memenuhi salah satu syarat dalam mencapai gelar Sarjana Komputer(S.Kom) jurusan Informatika Universitas Multimedia Nusantara. Saya menyadari bahwa pengerjaan laporan ini tidak akan bisa selesai jika tidak mendapatkan bimbingan dari berbagai pihak. Oleh karena itu saya mengucapkan terima kasih kepada:

1. Bapak Dr. Ir. Andrey Andoko, M.Sc., selaku Rektor Universitas Multimedia Nusantara.
2. Bapak Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Dekan Fakultas Teknik dan Informatika Universitas Multimedia Nusantara.
3. Bapak Arya Wicaksana, S.Kom., M.Eng.Sc., OCA, selaku Ketua Program Studi Informatika Universitas Multimedia Nusantara.
4. Bapak David Agustriawan, S.Kom., M.Sc., Ph.D., sebagai Pembimbing pertama yang telah memberikan bimbingan, arahan, dan motivasi atas terselesainya tugas akhir ini.
5. Keluarga dan teman-teman saya yang telah memberikan bantuan dukungan material dan moral, sehingga penulis dapat menyelesaikan tugas akhir ini.
6. Biwa Kawaii Putri Kato yang telah memberikan dukungan kepada penulis dalam menghadapi permasalahan.

Semoga tugas akhir ini bermanfaat bagi mahasiswa dan masyarakat yang ingin meneliti lebih lanjut tentang topik tugas akhir ini.

Tangerang, 1 Juli 2025



Agil Wira Pratama

**ANALISIS SENTIMEN PROGRAM MAKAN BERGIZI
GRATIS DI MEDIA SOSIAL X DENGAN
STACKING NAIVE BAYES
SVM DAN RF**

Agil Wira Pratama

ABSTRAK

Program Makan Bergizi Gratis yang diusung oleh pemerintah menimbulkan beragam opini publik di media sosial, khususnya Twitter. Penelitian ini bertujuan untuk menganalisis sentimen masyarakat terhadap program tersebut menggunakan pendekatan *Stacking Ensemble* yang menggabungkan tiga algoritma klasifikasi: Naive Bayes, Support Vector Machine (SVM), dan Random Forest. Pengumpulan data dilakukan melalui Twitter menggunakan kata kunci relevan, lalu diproses melalui tahap praproses teks dan representasi fitur dengan metode TF-IDF. Untuk mengatasi ketidakseimbangan data, digunakan teknik Synthetic Minority Oversampling Technique (SMOTE). Evaluasi model dilakukan menggunakan metrik akurasi, presisi, recall, dan F1-score. Hasil pengujian menunjukkan bahwa model *Stacking* menghasilkan performa lebih baik dibandingkan model dasar individu, dengan akurasi tertinggi mencapai 92,98% pada skenario pembagian data 90:10. Temuan ini menunjukkan bahwa pendekatan *ensemble* dapat meningkatkan efektivitas klasifikasi sentimen dalam studi opini publik.

Kata Kunci: Analisis Sentimen, Makan Bergizi Gratis, Naive Bayes, Random Forest, Stacking, SVM

U M N
U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

**SENTIMENT ANALYSIS OF THE NUTRITIOUS MEAL PROGRAM ON
SOCIAL MEDIA X USING STACKING NAIVE BAYES, SVM, RF**

Agil Wira Pratama

ABSTRACT

The Free Nutritious Meal Program initiated by the government has generated diverse public opinions on social media, particularly Twitter. This study aims to analyze public sentiment toward the program using a Stacking Ensemble approach that combines three classification algorithms: Naive Bayes, Support Vector Machine (SVM), and Random Forest. Data was collected from Twitter using relevant keywords, followed by preprocessing and feature representation using the TF-IDF method. To handle data imbalance, the Synthetic Minority Oversampling Technique (SMOTE) was applied. The models were evaluated using accuracy, precision, recall, and F1-score metrics. The results show that the Stacking model outperformed individual base classifiers, achieving the highest accuracy of 92,98% under the 90:10 data split. These findings indicate that ensemble approaches can enhance sentiment classification performance in public opinion analysis.

Keywords: *Free Nutritious Meals, Naive Bayes, Random Forest, Sentiment Analysis, Stacking, SVM*



DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERSETUJUAN	iv
HALAMAN PERSEMBAHAN/MOTO	v
KATA PENGANTAR	vi
ABSTRAK	vii
ABSTRACT	viii
DAFTAR ISI	ix
DAFTAR TABEL	xi
DAFTAR GAMBAR	xii
DAFTAR KODE	xiii
DAFTAR RUMUS	xiv
DAFTAR LAMPIRAN	xv
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	2
1.3 Batasan Permasalahan	2
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.6 Sistematika Penulisan	4
BAB 2 LANDASAN TEORI	5
2.1 Analisis Sentimen	5
2.2 Media Sosial X (Twitter)	5
2.3 Tweet Harvest	5
2.4 TF-IDF (Term Frequency - Inverse Document Frequency)	6
2.4.1 Term Frequency (TF)	6
2.4.2 Inverse Document Frequency (IDF)	6
2.5 SMOTE	7
2.6 Algoritma Klasifikasi	7
2.6.1 Algoritma Naive Bayes	8
2.6.2 Chi-Square	9
2.6.3 Algoritma Support Vector Machine	10
2.6.4 Decision Tree	12
2.6.5 Random Forest	14
2.7 Hyperparameter Tuning	14
2.8 Ensemble Learning	14
2.8.1 Stacking	15
2.9 Confusion Matirx	15
2.9.1 Accuracy	16
2.9.2 Precision	16
2.9.3 Recall	17
2.9.4 F1-Score	17
BAB 3 METODOLOGI PENELITIAN	18
3.1 Metodologi Penelitian	18
3.2 Alur Metodologi Penelitian	19
3.2.1 Studi Literatur	20
3.2.2 Crawling Data	20

3.2.3	Preprocessing & Labeling Data	21
3.2.4	Feature Engineering	28
3.2.5	Perancangan Model	29
3.2.6	Implementasi dan Pengujian	39
BAB 4	HASIL DAN DISKUSI	42
4.1	Spesifikasi Sistem	42
4.2	Implementasi Sistem	42
4.2.1	Crawling Data	42
4.2.2	Preprocessing Data	43
4.2.3	Labelling Data	50
4.2.4	Feature Engineering	51
4.2.5	Balancing Data (SMOTE)	51
4.2.6	Feature Selection	52
4.2.7	Hyperparameter Tuning	52
4.2.8	Stacking	54
4.2.9	Evaluasi Best Model	55
4.3	Pengujian Baseline Model dan Ensemble Model	55
4.3.1	Pengujian Ensemble Learning Stacking Chi-Square	60
4.4	Kesimpulan Pengujian Model	61
BAB 5	SIMPULAN DAN SARAN	63
5.1	Simpulan	63
5.2	Saran	63
DAFTAR PUSTAKA		65



DAFTAR TABEL

Tabel 4.1	Perbandingan sebelum dan sesudah case folding	44
Tabel 4.2	Perbandingan sebelum dan sesudah cleaning teks	45
Tabel 4.3	Perbandingan sebelum dan tokenisasi data	46
Tabel 4.4	Perbandingan sebelum dan sesudah stopword removed	47
Tabel 4.5	Perbandingan sebelum dan sesudah stemming	48
Tabel 4.6	Perbandingan sebelum dan sesudah normalisasi data	49
Tabel 4.7	10 kata teratas berdasarkan skor tf-idf	51
Tabel 4.8	Hasil pengujian SelectKBest terhadap akurasi model	52
Tabel 4.9	Hasil hyperparameter tuning svm	53
Tabel 4.10	Hasil hyperparameter tuning random forest	53
Tabel 4.11	Hasil hyperparameter tuning multinomial naive bayes	54
Tabel 4.12	Hasil evaluasi model stacking dengan SelectKBest (k='all')	54
Tabel 4.13	Hasil evaluasi training SVM (RBF kernel)	56
Tabel 4.14	Hasil evaluasi testing SVM (RBF kernel)	56
Tabel 4.15	Hasil evaluasi training base model Multinomial Naive Bayes	57
Tabel 4.16	Hasil evaluasi testing base model Multinomial Naive Bayes	57
Tabel 4.17	Hasil evaluasi training Random Forest	57
Tabel 4.18	Hasil evaluasi testing Random Forest	57
Tabel 4.19	Hasil evaluasi training SVM (RBF kernel)	58
Tabel 4.20	Hasil evaluasi testing SVM (RBF kernel)	58
Tabel 4.21	Hasil evaluasi training Multinomial Naive Bayes (90:10)	59
Tabel 4.22	Hasil evaluasi testing Multinomial Naive Bayes (90:10)	59
Tabel 4.23	Hasil evaluasi training Random Forest (90:10)	59
Tabel 4.24	Hasil evaluasi testing Random Forest (90:10)	59
Tabel 4.25	Hasil evaluasi training Stacking Classifier (tuning, 90:10)	60
Tabel 4.26	Hasil evaluasi testing Stacking Classifier (tuning, 90:10)	60
Tabel 4.27	Hasil evaluasi training model Stacking (Chi2, 90:10)	61
Tabel 4.28	Hasil evaluasi testing model Stacking (Chi2, 90:10)	61
Tabel 4.29	Perbandingan Akurasi Model: Base Model dan Ensemble Learning Before Tuning dengan Chi2-Stacked	61
Tabel 4.30	Perbandingan Akurasi Model: Base Model dan Ensemble Learning (After Tuning) dengan Chi2-Stacked	62

UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR GAMBAR

Gambar 2.1	Ilustrasi smote oversampling	7
Gambar 2.2	Ilustrasi svm	10
Gambar 2.3	Confusion matrix table	16
Gambar 3.1	<i>Pipeline metodologi penelitian</i>	19
Gambar 3.2	<i>Flowchart</i> pengumpulan data dengan metode <i>crawling</i> . .	20
Gambar 3.3	<i>Flowchart preporcessing data</i>	22
Gambar 3.4	<i>Flowchart labeling data</i> dengan <i>indobert</i>	26
Gambar 3.5	<i>Flowchart tf-idf</i>	29
Gambar 3.6	Flowchart perancangan model pengujian	30
Gambar 3.7	Flowchart <i>feature selection</i> menggunakan SelectKBest . .	31
Gambar 3.8	Flowchart <i>pemodelan baseling</i> MNB, SVM, dan RF	32
Gambar 3.9	Flowchart Hyperparameter Tuning	34
Gambar 3.10	<i>Flowchart</i> model stacking	38
Gambar 4.1	Hasil crawling data	43
Gambar 4.2	Preprocessing Case Folding	44
Gambar 4.3	Proses <i>preprocessing case folding</i>	45
Gambar 4.4	Preprocessing tokenisasi data	46
Gambar 4.5	Preprocessing <i>stopword removed</i>	47
Gambar 4.6	Preprocessing <i>stemming data</i>	48
Gambar 4.7	Preprocessing normalisasi data	49
Gambar 4.8	Hasil labeling data	50
Gambar 4.9	Jumlah Labeling Data	50
Gambar 4.10	Jumlah distribusi data setelah SMOTE	52
Gambar 4.11	Confusion Matrix Model Stacking (SelectKBest, k='all') .	55



DAFTAR KODE

Kode 3.1	Konfigurasi token dan instalasi tweet harvest	21
Kode 3.2	Menentukan keyword dan menyimpan hasil crawling	21
Kode 3.3	Menampilkan data tweet ke dalam tabel	22
Kode 3.4	Case Folding: Mengubah teks menjadi huruf kecil	23
Kode 3.5	Cleaning: Membersihkan simbol pada data	23
Kode 3.6	Tokenisasi: Memisahkan kata-kata dalam teks	24
Kode 3.7	Stopword Removal: Menghapus kata umum yang tidak penting .	24
Kode 3.8	Stemming: Mengubah kata ke bentuk dasar	25
Kode 3.9	Normalisasi: Menghapus kata tidak baku atau slang	25
Kode 3.10	Labeling otomatis data menggunakan model indoBERT	27
Kode 3.11	pengaplikasian TF-IDF	29
Kode 3.12	Pengaplikasian splitting data	30
Kode 3.13	Pengaplikasian SMOTE	30
Kode 3.14	Implementasi kode pengujian feature seletiton	31
Kode 3.15	Implementasi kode pengujian SVM	32
Kode 3.16	Implementasi kode pengujian Random Forest	33
Kode 3.17	Implementasi kode pengujian Multinomial Naive Bayes	33
Kode 3.18	Implementasi hyperparameter tuning SVM	34
Kode 3.19	Implementasi hyperparameter tuning random forest	35
Kode 3.20	Implementasi hyperparameter tuning naive bayes	36
Kode 3.21	Implementasi Kode Penggunaan Stacking Classifier	38



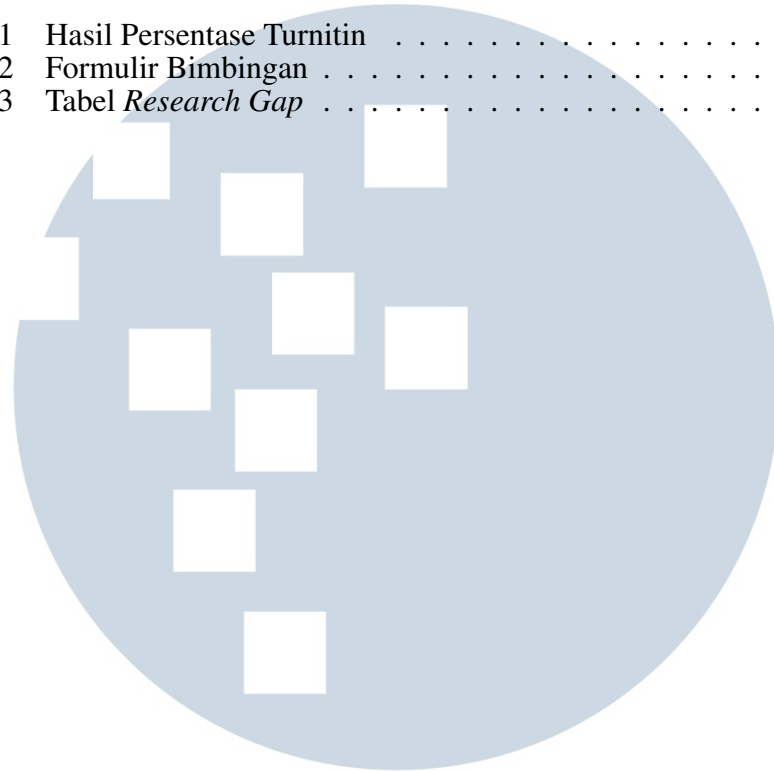
DAFTAR RUMUS

Rumus 2.1	<i>TF-IDF (Term Frequency - Inverse Document Frequency)</i> . . .	6
Rumus 2.2	<i>Term Frequency (TF)</i>	6
Rumus 2.3	<i>Inverse Document Frequency (IDF)</i>	7
Rumus 2.4	<i>Teorema Bayes</i>	8
Rumus 2.5	<i>Multinomial Naive Bayes</i>	9
Rumus 2.6	<i>Multinomial Naive Bayes dengan TF-IDF</i>	9
Rumus 2.7	<i>Chi-Square Score</i>	10
Rumus 2.8	<i>Prinsip Dasar SVM</i>	11
Rumus 2.9	<i>SVM Fungsi Objektif</i>	11
Rumus 2.10	<i>SVM Fungsi Objektif dengan Kendala</i>	11
Rumus 2.11	<i>SVM Linear Kernel</i>	12
Rumus 2.12	<i>SVM RBF Kernel</i>	12
Rumus 2.13	<i>Decision Tree Entropy</i>	13
Rumus 2.14	<i>Decision Tree Gain</i>	13
Rumus 2.15	<i>Confusion Matrix Accuracy</i>	16
Rumus 2.16	<i>Confusion Matrix Precision</i>	16
Rumus 2.17	<i>Confusion Matrix Recall</i>	17
Rumus 2.18	<i>Confusion Matrix F1-Score</i>	17



DAFTAR LAMPIRAN

Lampiran 1	Hasil Persentase Turnitin	69
Lampiran 2	Formulir Bimbingan	70
Lampiran 3	Tabel <i>Research Gap</i>	73



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA