



**Perbandingan SVM dan Hybrid
IndoBERTweet–SVM untuk Analisis
Sentimen Program MBG di Media Sosial**

**TUGAS AKHIR
SKRIPSI**

**UNIVERSITAS
MERCU BUANA**
**ALDI RAHMAN
HERMAWAN
41822010004**

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2026**



**Perbandingan SVM dan Hybrid
IndoBERTweet–SVM untuk Analisis
Sentimen Program MBG di Media Sosial**

**TUGAS AKHIR
SKRIPSI**

**ALDI RAHMAN
HERMAWAN
41822010004**

**UNIVERSITAS
MERCU BUANA**

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

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2026**

HALAMAN PERNYATAAN KARYA SENDIRI

Saya yang bertanda tangan di bawah ini:

Nama : Aldi Rahman Hermawan
NIM : 41822010004
Program Studi : Sistem Informasi

Menyatakan dengan sebenar-benarnya bahwa Tugas Akhir berjudul:
“Perbandingan SVM dan Hybrid IndoBERTweet-SVM untuk Analisis Sentimen Program MBG di Media Sosial”

adalah hasil karya saya sendiri, tidak mengandung unsur plagiarisme, pelanggaran hak cipta, atau konten ilegal dalam bentuk apapun dan tidak melanggar hukum atau hak pihak manapun.

Apabila di kemudian hari ditemukan pelanggaran terhadap pernyataan ini, saya bersedia menanggung seluruh konsekuensi hukum dan membebaskan Universitas Mercu Buana dari segala bentuk tuntutan hukum dan saya siap mendapatkan sanksi akademis yang berlaku di Universitas Mercu Buana.

Demikian pernyataan ini saya buat dengan sebenar-benarnya untuk digunakan sebagaimana mestinya.

Jakarta, 17 Agustus 2026



Aldi Rahman Hermawan

UNIVERSITAS
MERCU BUANA

PERNYATAAN SIMILARITY CHECK
(SIMILARITY CHECK STATEMENT)

Saya yang bertanda tangan dibawah ini menyatakan, bahwa karya ilmiah yang ditulis oleh
The undersigned, hereby declare that the scientific work written by:

Nama /Name : Aldi Rahman Hermawan
NIM /Student ID Number : 41822010004
Program Studi /Study Program : Sistem Informasi

Dengan Judul Tugas Akhir

The title:

“Perbandingan SVM dan Hybrid IndoBERTweet-SVM untuk Analisis Sentimen Program MBG di Media Sosial”

telah dilakukan pengecekan *similarity* dengan sistem Turnitin pada tanggal:

Similarity checks have been carried out with the Turnitin system on the date:

18 Agustus 2026

dengan nilai persentase sebesar :

with a percentage value of:

23%

dinyatakan memenuhi standar sesuai dengan ketentuan berlaku di Fakultas Ilmu Komputer Universitas Mercu Buana. *declared to meet standards in accordance with applicable regulations at the Faculty of Computer Science Universitas Mercu Buana.*

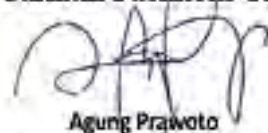
File hasil cek similarity turnitin:

Turnitin similarity report file

https://drive.google.com/file/d/1Ems5kLJp-muGfivXwvaUCzsMGyPD0rSNil/view?usp=drive_link



Jakarta, 18 Agustus 2026
Admin Turnitin Fasilkom UMB



Agung Prawoto, S.Kom., B.Sc
NIK : 322970503

HALAMAN PENGESAHAN

Laporan Skripsi ini diajukan oleh:

Nama : Aldi Rahman Hermawan
NIM : 41822010004
Program Studi : Sistem Informasi
Judul Laporan Skripsi : Perbandingan SVM dan Hybrid IndoBERTweet–
SVM untuk Analisis Sentimen Program MBG di
Media Sosial

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 Sistem Informasi, Fakultas Ilmu Komputer Universitas Mercu Buana.

Disahkan oleh:

Pembimbing,



Wawan Gunawan, S.Kom., M.T., M.Kom.
NIDN/NUPTK: 0424108104

UNIVERSITAS

MERCU BUANA

Jakarta, 18 Agustus 2026

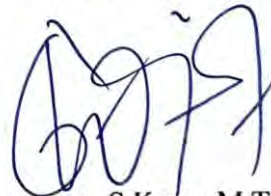
Mengetahui,

Dekan Fakultas Ilmu Komputer



Dr. Bambang Jokonowo, S.Si., M.T.I.
NIDN: 0320037002

Ketua Program Studi
Sistem Informasi



Wawan Gunawan, S.Kom., M.T., M.Kom.
NIDN: 0424108104

KATA PENGANTAR

Puji syukur kepada Allah Yang Maha Kuasa atas rahmat dan berkah-Nya, saya telah berhasil menyelesaikan Tugas Akhir ini. Laporan ini disusun sebagai salah satu persyaratan untuk mengikuti sidang Tugas Akhir. Saya menyadari bahwa penyelesaian Tugas Akhir ini tidak akan mungkin terwujud tanpa dukungan dan bimbingan dari berbagai pihak, mulai dari masa saya sebagai mahasiswa hingga proses penulisan. Oleh karena itu, saya ingin mengucapkan terima kasih kepada semua pihak yang telah membantu saya:

1. Bapak Prof. Dr. Andi Adriansyah, M.Eng selaku Rektor Universitas Mercu Buana;
2. Bapak Dr. Bambang Jokonowo, S.Si., M.T.I selaku Dekan Fakultas Ilmu Komputer;
3. Bapak Wawan Gunawan, S.Kom., MT., M.Kom selaku Ketua Program Studi Sistem Informasi sekaligus Dosen Pembimbing atas bimbingan dan dukungan yang beliau berikan kepada peneliti;
4. Orang tua, keluarga, serta teman-teman yang senantiasa memberikan doa, dukungan, dan motivasi kepada penulis.

Akhir kata, penulis berharap semoga Tugas Akhir ini dapat memberikan manfaat bagi pengembangan ilmu pengetahuan, khususnya di bidang teknologi informasi, serta dapat menjadi referensi bagi penelitian selanjutnya.

Jakarta, 24 Juli 2026



Aldi Rahman Hermawan

**HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI
TUGAS AKHIR DI REPOSITORI UMB**

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

Nama : Aldi Rahman Hermawan
NIM : 41822010004
Fakultas/Program Studi : Sistem Informasi
Judul Tugas Akhir : Perbandingan SVM dan Hybrid IndoBERTweet-SVM untuk Analisis Sentimen Program MBG di Media Sosial

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 Tugas Akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta. Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 18 Agustus 2026
Yang menyatakan,

UNIVERSITAS
MERCU BUANA

(Aldi Rahman Hermawan)



**PERBANDINGAN SVM DAN HYBRID INDOBERTWEET–SVM UNTUK ANALISIS
SENTIMEN PROGRAM MBG DI MEDIA SOSIAL
ALDI RAHMAN HERMAWAN**

ABSTRAK

Program Makan Bergizi Gratis adalah salah satu kebijakan pemerintah Indonesia yang banyak memperoleh perhatian masyarakat di media sosial. Penelitian ini bertujuan menganalisis sentimen publik terhadap Program MBG serta membandingkan kinerja model Support Vector Machine (SVM) berbasis Term Frequency–Inverse Document Frequency (TF-IDF) dengan model Hybrid IndoBERTweet–SVM. Sentimen publik dalam penelitian ini didefinisikan sebagai kecenderungan opini masyarakat yang dikelompokkan ke dalam kelas positif, netral, dan negatif, sedangkan kinerja model diukur menggunakan accuracy, precision, recall, dan F1-score. Sampel penelitian terdiri atas 2.335 data teks yang berasal dari Twitter/X dan komentar YouTube, dengan distribusi 545 data positif, 569 data netral, dan 1.221 data negatif. Data dipilih menggunakan teknik purposive sampling berdasarkan kesesuaian isi unggahan dan komentar dengan topik Program MBG. Analisis dilakukan menggunakan pendekatan kuantitatif komparatif melalui tahapan pengumpulan data, pembersihan teks, normalisasi, pelabelan, pembagian data, pelatihan model, optimasi parameter, dan evaluasi. Pada model pertama, teks direpresentasikan menggunakan TF-IDF dan diklasifikasikan dengan SVM, sedangkan pada model kedua, embedding kontekstual IndoBERTweet digunakan sebagai fitur masukan bagi SVM. Hasil penelitian menunjukkan bahwa model Hybrid IndoBERTweet–SVM memperoleh test accuracy sebesar 72,59% dan macro F1-score sebesar 70,67%, sedangkan model SVM berbasis TF-IDF memperoleh test accuracy sebesar 61,88% dan macro F1-score sebesar 59,58%. Model hybrid juga menghasilkan nilai F1-score yang lebih tinggi pada seluruh kelas sentimen, meskipun kelas netral tetap menjadi kategori yang paling sulit diklasifikasikan. Berdasarkan hasil tersebut, embedding kontekstual IndoBERTweet lebih efektif dibandingkan pembobotan TF-IDF dalam merepresentasikan bahasa media sosial Indonesia dan meningkatkan kemampuan SVM dalam membedakan sentimen positif, netral, dan negatif.

Kata Kunci : Analisis Sentimen, Media Sosial, SVM, IndoBERTweet.

COMPARISON OF SVM AND HYBRID INDOBERTWEET–SVM FOR SENTIMENT ANALYSIS OF THE FREE NUTRITIOUS MEAL PROGRAM ON SOCIAL MEDIA

ALDI RAHMAN HERMAWAN

ABSTRACT

The Free Nutritious Meals Program is an Indonesian government policy that has attracted considerable public attention on social media. This study aims to analyze public sentiment toward the MBG Program and compare the performance of a Support Vector Machine (SVM) using Term Frequency–Inverse Document Frequency (TF-IDF) with a Hybrid IndoBERTweet–SVM model. Public sentiment is defined as the tendency of public opinion classified into positive, neutral, and negative categories, while model performance is assessed using accuracy, precision, recall, and F1-score. The research sample consists of 2,335 text data collected from Twitter/X and YouTube comments, comprising 545 positive, 569 neutral, and 1,221 negative instances. The data were selected using purposive sampling based on the relevance of posts and comments to the MBG Program. A quantitative comparative approach was employed through data collection, text cleaning, normalization, sentiment labeling, data splitting, model training, parameter optimization, and evaluation. In the first model, texts were represented using TF-IDF and classified with SVM. In the second model, contextual embeddings generated by IndoBERTweet were used as input features for SVM. The results show that the Hybrid IndoBERTweet–SVM achieved a test accuracy of 72.59% and a macro F1-score of 70.67%, whereas the TF-IDF-based SVM obtained a test accuracy of 61.88% and a macro F1-score of 59.58%. The hybrid model also produced higher F1-scores across all sentiment classes, although the neutral class remained the most difficult to classify. These findings indicate that IndoBERTweet contextual embeddings are more effective than TF-IDF weighting in representing Indonesian social media language and improving SVM performance in distinguishing positive, neutral, and negative sentiments.

Keywords: Sentiment Analysis, Social Media, SVM, Transformer.

DAFTAR ISI

HALAMAN SAMPUL	0
HALAMAN JUDUL.....	i
HALAMAN PERNYATAAN KARYA SENDIRI.....	ii
HALAMAN SURAT KETERANGAN UJI TURNITIN	iii
HALAMAN PENGESAHAN.....	iv
KATA PENGANTAR.....	v
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR DI REPOSITORI UMB.....	vi
ABSTRAK.....	vii
ABSTRACT.....	viii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xi
DAFTAR GAMBAR	xii
DAFTAR LAMPIRAN.....	xiii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah	4
1.3 Tujuan Penelitian	4
1.4 Batasan Masalah.....	4
1.5 Sistematika Penulisan.....	5
BAB II TINJAUAN PUSTAKA.....	6
2.1 Teori Utama.....	6
2.1.1 Support Vector Machine (SVM).....	6
2.1.2 IndoBERTweet	7
2.1.3 Analisis Sentimen	8
2.2 Teori Pendukung	8
2.2.1 Program Makan Bergizi Gratis	8
2.2.2 Twitter (X).....	8
2.2.3 YouTube	9
2.3 Penelitian Terdahulu	9
2.4 Gap Penelitian	22
BAB III METODE PENELITIAN.....	23
3.1 Pendekatan Penelitian	23
3.2 Desain Penelitian.....	24
3.3 Instrumen Penelitian.....	25
3.4 Tahapan Penelitian.....	26
3.4.1 Studi Literatur.....	27
3.4.2 Pengumpulan Data	27
3.4.3 Preprocessing Data.....	29
3.4.4 Pelabelan Data	32
3.4.5 Pembagian Data	33
3.4.6 Ekstraksi Fitur dan Pelatihan Model	34
3.4.7 Evaluasi Model	35
3.5 Teknik Pengumpulan Data.....	35
3.6 Kesetaraan Twitter/X dan YouTube sebagai Sumber Data	36
BAB IV PEMBAHASAN.....	38
4.1 Persiapan Data	38
4.2 Pre-Processing Data	41
4.3 Pelabelan Data.....	50
4.4 Pembagian Data (Splitting Data)	54
4.5 Permodelan.....	55
4.5.1 Implementasi Model SVM berbasis TF-IDF	55
4.5.2 Implementasi Model Hybrid IndoBERTweet-SVM.....	59
4.6 Evaluasi Model.....	65
BAB V KESIMPULAN	71

5.1	Kesimpulan.....	71
5.2	Saran.....	72
DAFTAR PUSTAKA.....		74
LAMPIRAN.....		76



DAFTAR TABEL

Table 2. 1 Penelitian Terdahulu.....	9
Table 3. 1 Spesifikasi Hardware.....	25
Table 3. 2 Versi Software.....	25
Table 3. 3 Sample Data.....	29
Table 3. 4 Hasil Teks Normalisasi.....	30
Table 4. 1 Konfigurasi Twitter Crawler.....	38
Table 4. 2 Contoh Beberapa Sampel Dataset Dari Twitter.....	39
Table 4. 3 Contoh Beberapa Sampel Dataset Dari Youtube.....	40
Table 4. 4 Perbandingan Karakteristik Data Twitter/X Dan Youtube.....	40
Table 4. 5 Contoh Pre-Processing Tahap Case Folding.....	42
Table 4. 6 Contoh Pre-Processing Tahap Cleaning.....	43
Table 4. 7 Kamus Normalisasi Yang Digunakan.....	44
Table 4. 8 Contoh Pre-Processing Tahap Text Normalisasi.....	45
Table 4. 9 Contoh Pre-Processing Tahap Tokenization.....	46
Table 4. 10 Contoh Pre-Processing Tahap Stopword Removal.....	47
Table 4. 11 Contoh Pre-Processing Tahap Stemming.....	49
Table 4. 12 Contoh Perbandingan Antar Label Llm Dan Label Anotator.....	50
Table 4. 13 Distribusi Persentase Label Llm Dan Label Anotator.....	52
Table 4. 14 Jumlah Dataset Yang Relevan.....	53
Table 4. 15 Pembagian Dataset.....	55
Table 4. 16 Contoh Hasil Pembobotan Tf-Idf.....	56
Table 4. 17 Hasil Uji Parameter ϵ	57
Table 4. 18 Classification Report Svm Tf-Idf.....	58
Table 4. 19 Hasil Fine Tuning Indobertweet Per Fold.....	60
Table 4. 20 Hasil Uji Parameter C Untuk Model Hybrid.....	62
Table 4. 21 Classification Report Hybrid Indobertweet-Svm.....	63
Table 4. 22 Perbandingan Hasil Evaluasi Model.....	66
Table 4. 23 Perbandingan F1-Score Setiap Kelas.....	67

UNIVERSITAS
MERCU BUANA

DAFTAR GAMBAR

Gambar 2. 1 Support Vector Machine	7
Gambar 3. 1 Desain Penelitian	25
Gambar 3. 2 Tahapan Penelitian	27
Gambar 3. 3 Tahapan Crawling Data Platform Twitter/X.....	28
Gambar 3. 4 Tahapan Crawling Data Platform Youtube.....	28
Gambar 4. 1 Distribusi Sentimen Komentar Mbg.....	53
Gambar 4. 2 Perbandingan Label Relevan Dan Tidak Relevan.....	54
Gambar 4. 3 Confusion Matrix Svm+Tf-Idf.....	59
Gambar 4. 4 Grafik Training Dan Validation Loss Model Indobertweet.....	61
Gambar 4. 5 Confusion Matrix Hybrid Indobertweet-Svm	65
Gambar 4. 6 Grafik Perbandingan Metrik Evaluasi Model	67
Gambar 4. 7 Grafik Perbandingan F1-Score Per Kelas	68
Gambar 4. 8 Perbandingan Confusion Matrix Svm Tf-Idf Dan Hybrid Indobertweet-Svm.....	69



DAFTAR LAMPIRAN

Lampiran 1 Kartu Asistensi.....	76
Lampiran 2 Curriculum Vitae.....	77
Lampiran 3 Bukti Bnsp	78
Lampiran 4 Lembar Pernyataan Hasil Cek Uji Plagiasi	79

