

**ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI
JOGJAKARTA PADA GOOGLE PLAY STORE
MENGUNAKAN ALGORITMA SUPPORT VECTOR
MACHINE**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi SI Informatika



disusun oleh

FERNANDO EGA HOTAMA

22.11.4904

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2026**

**ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI
JOGJAKARTA PADA GOOGLE PLAY STORE
MENGUNAKAN ALGORITMA SUPPORT VECTOR
MACHINE**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi SI Informatika



disusun oleh

FERNANDO EGA HOTAMA

22.11.4904

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2026

HALAMAN PERSETUJUAN

SKRIPSI

**ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI JOGJAKARTA
PADA GOOGLE PLAY STORE MENGGUNAKAN ALGORITMA
SUPPORT VECTOR MACHINE**

yang disusun dan diajukan oleh

Fernando Ega Hotama

22.11.4904

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 27 Januari 2026

Dosen Pembimbing,



Rumini, S.Kom., M.Kom.
NIK. 190302246

HALAMAN PENGESAHAN

SKRIPSI

**ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI JOGJAKARTA
PADA GOOGLE PLAY STORE MENGGUNAKAN ALGORITMA
SUPPORT VECTOR MACHINE**

yang disusun dan diajukan oleh

Fernando Ega Hotama

22.11.4904

Telah dipertahankan di depan Dewan Penguji
pada tanggal 27 Januari 2026

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Windha Mega Pradnya Dhuhiita, S.Kom., M.Kom.
NIK. 190302185



Melany Mustika Dewi, S.Kom., M.Kom.
NIK. 190302455



Rumini, S.Kom., M.Kom.
NIK. 190302246



Skrripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 27 Januari 2026

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusriani, M.Kom.
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Fernando Ega Hotama
NIM : 22.11.4904

Menyatakan bahwa Skripsi dengan judul berikut:

Analisis Sentimen Ulasan Pengguna Aplikasi Jogjakita Pada Google Play Store Menggunakan Algoritma Support Vector Machine

Dosen Pembimbing : Rumini, S.Kom., M.Kom.

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 27 Januari 2026

Yang Menyatakan,



Fernando Ega Hotama

HALAMAN PERSEMBAHAN

Puji dan syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa atas rahmat dan karunia-Nya yang senantiasa memberikan kekuatan, ketekunan, dan hikmat dalam menyelesaikan skripsi ini. Karya ilmiah ini merupakan hasil dari proses panjang yang penuh pembelajaran, perjuangan, dan doa yang tulus. Skripsi ini penulis persembahkan kepada kedua orang tua tercinta, sebagai ungkapan kasih dan penghargaan atas segala doa, dukungan, dan pengorbanan yang tiada henti. Persembahan ini juga ditujukan kepada dosen pembimbing dan seluruh pengajar di Universitas Amikom Yogyakarta, yang telah menjadi bagian penting dalam perjalanan akademik penulis. Selain itu, karya ini penulis persembahkan kepada teman-teman dan saudara-saudara yang selalu memberikan semangat dan inspirasi, serta kepada diri sendiri, sebagai wujud apresiasi atas usaha, ketekunan, dan keyakinan dalam menyelesaikan setiap proses hingga akhir. Semoga karya ini menjadi langkah kecil yang bermakna dan dapat memberi manfaat bagi pengembangan ilmu pengetahuan di masa mendatang.

KATA PENGANTAR

Puji dan syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa atas rahmat, kasih, dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul **“Analisis Sentimen Ulasan Pengguna Aplikasi JogjaKita Menggunakan Algoritma Support Vector Machine (SVM)”** dengan baik. Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer pada Program Studi SI Informatika, Fakultas Ilmu Komputer, Universitas Amikom Yogyakarta.

Pada kesempatan ini penulis ingin menyampaikan ucapan terima kasih kepada:

1. Rumi, M.Kom., selaku dosen pembimbing skripsi yang dengan sabar telah memberikan arahan, masukan, serta waktu dalam membimbing penulis hingga skripsi ini dapat diselesaikan dengan baik.
2. Kedua orang tua tercinta, atas kasih, doa, dan dukungan sebagai motivasi dalam menyelesaikan studi.
3. Teman-teman kelas IF 06 Angkatan 2022, atas kebersamaan, semangat, dan dukungan yang diberikan selama masa perkuliahan dan proses penyusunan skripsi.
4. Teman-teman OMK Gereja Maria Assumpta Pakem yang senantiasa menghibur dan selalu memberi dukungan serta menjadi tempat berkeluh kesah.
5. Seluruh pihak yang tidak bisa diucapkan satu persatu yang telah membantu penulis dalam pengerjaan skripsi ini.

Penulis menyadari bahwa skripsi ini masih memiliki kekurangan. Oleh karena itu, kritik dan saran yang membangun demi penyempurnaan di masa mendatang.

Yogyakarta, 27 Januari 2026



Fernando Ega Hotama

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI.....	vii
DAFTAR TABEL.....	x
DAFTAR GAMBAR.....	xi
DAFTAR LAMPIRAN.....	xii
DAFTAR LAMBANG DAN SINGKATAN	xiii
DAFTAR ISTILAH	xiv
INTISARI	xv
<i>ABSTRACT</i>	xvi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	2
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	4
1.6 Sistematika Penulisan	4
BAB II TINJAUAN PUSTAKA	6
2.1 Studi Literatur	6

2.2	Dasar Teori.....	11
2.2.1	<i>Data Mining dan Machine Learning</i>	11
2.2.2	Analisis Sentimen.....	11
2.2.3	<i>Text Mining</i>	12
2.2.4	<i>Natural Language Processing (NLP)</i>	13
2.2.5	<i>Text Preprocessing</i>	13
2.2.6	Ekstraksi Fitur (TF-IDF).....	14
2.2.7	<i>Hyperparameter Tunning</i>	15
2.2.8	<i>Algoritma Support Vector Machine (SVM)</i>	16
2.2.9	Evaluasi Kinerja Model.....	19
2.2.10	Implementasi Sistem.....	21
BAB III METODE PENELITIAN		23
3.1	Objek Penelitian.....	23
3.2	Alur Penelitian	23
3.2.1	Pengumpulan Data (<i>web scraping</i>).....	24
3.2.2	<i>Preprocessing Data</i>	25
3.2.3	Labeling Data.....	31
3.2.4	<i>Split Data</i>	34
3.2.5	<i>Ekstraksi Fitur TF-IDF</i>	35
3.2.6	Optimasi Hyperparameter.....	40
3.2.7	<i>Pelatihan Model Support Vector Machine (SVM)</i>	40
3.2.8	Evaluasi Model.....	43
3.2.9	Implementasi Sistem Berbasis Streamlit.....	45
3.3	Alat dan Bahan.....	46
BAB IV HASIL DAN PEMBAHASAN		48
4.1	Pengumpulan Data	48
4.2	Preprocessing Teks	49
4.3	Pelabelan Sentimen	51
4.4	Pembagian Data	52
4.5	Ekstraksi Fitur TF-IDF	54

4.6	Optimasi Hyperparameter SVM	56
4.7	Pelatihan model SVM	57
4.7.1	Pelatihan Model Kernel Linear	57
4.7.2	Pelatihan Model Kernel RBF	58
4.8	Evaluasi Model	60
4.8.1	Evaluasi Model Kernel Linear	60
4.8.2	Evaluasi Model Kernel RBF	61
4.8.3	Evaluasi Hyperparameter Tunning (GridSearchCV)	62
4.8.4	Ringkasan Perbandingan Hasil Evaluasi Model SVM	64
4.9	Analisis WordCloud Sentimen	65
4.10	Implementasi Sistem	67
BAB V PENUTUP		70
5.1	Kesimpulan	70
5.2	Saran	71
DAFTAR PUSTAKA		72
LAMPIRAN		76
1.	Source code web scraping	76
2.	Source code Analisis sentimen menggunakan SVM	77
3.	Link Implementasi Sistem	82

DAFTAR TABEL

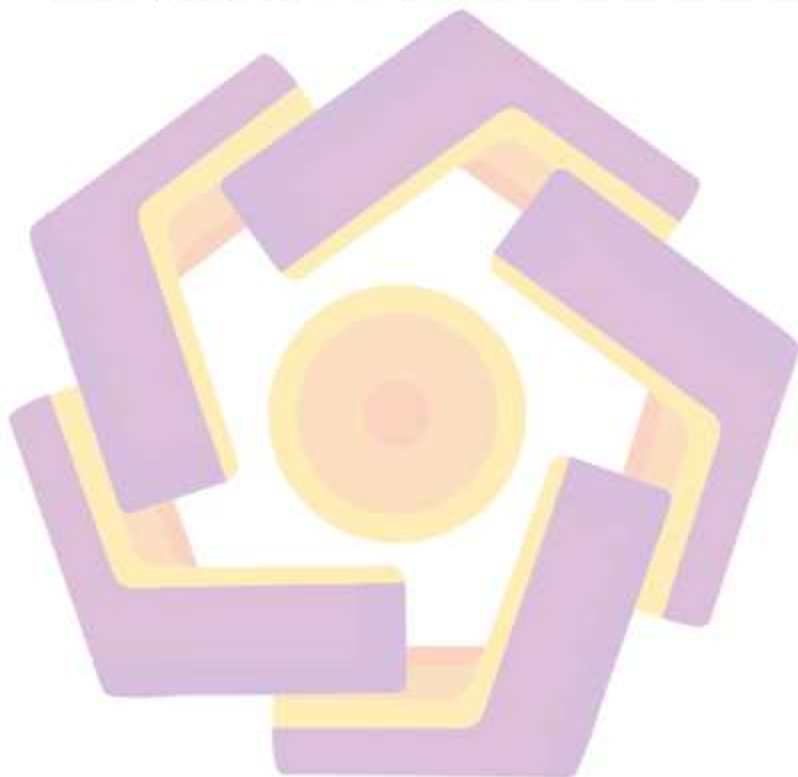
Tabel 2.1 Keaslian Penelitian	9
Tabel 3.1 Hasil case folding ulasan pengguna.....	26
Tabel 3.2 Hasil cleaning ulasan pengguna.....	27
Tabel 3.3 Hasil tahap tokenizing.....	28
Tabel 3.4 Hasil stopword removal.....	30
Tabel 3.5 Hasil tahap stemming.....	31
Tabel 3.6 Hasil labeling.....	32
Tabel 3.7 Hasil perhitungan ulasan positif.....	33
Tabel 3.8 Hasil perhitungan ulasan negatif.....	33
Tabel 3.9 Hasil pembagian data.....	35
Tabel 3.10 Distribusi Label data training.....	35
Tabel 3.11 Term Frequency (TF).....	36
Tabel 3.12 Document Frequency (DF).....	37
Tabel 3.13 Inverse Document Frequency.....	38
Tabel 3.14 Bobot TF-IDF.....	39
Tabel 3.15 Bobot TF-IDF lima dokumen ulasan.....	41
Tabel 3.16 Confusion Matrix.....	44
Tabel 3.17 Spesifikasi perangkat keras.....	46
Tabel 3.18 Perangkat lunak yang digunakan.....	47
Tabel 4.1 Split Data Training dan Testing.....	53
Tabel 4.2 Perbandingan Hasil Evaluasi Model Berdasarkan Data Uji.....	64

DAFTAR GAMBAR

Gambar 2.1 Ilustrasi hyperplane pada algoritma SVM [19].....	17
Gambar 2.2 True dan False positif dan negatif [29].	21
Gambar 3.2 Diagram Alur Penelitian.	24
Gambar 3.3 Hasil web scraping menggunakan google-play-scraper.....	25
Gambar 4.2 Dataset mentah.	49
Gambar 4.3 Hasil Preprocessing Teks.	50
Gambar 4.4 Hasil labeling.	51
Gambar 4.5 Distribusi Label Sentimen.	52
Gambar 4.6 Distribusi Data Training.	53
Gambar 4.7 Distribusi Data Testing.	54
Gambar 4.8 Ukuran Matriks TF-IDF Data Latih.	55
Gambar 4.9 Hasil Ekstraksi fitur TF-IDF (Unigram-Bigram).	55
Gambar 4.10 Hasil Optimasi Parameter SVM Menggunakan GridSearchCV.	57
Gambar 4.11 Evaluasi Training SVM Linear.	58
Gambar 4.12 Confusion Matrix Training SVM Kernel Linear.	58
Gambar 4.13 Evaluasi Training SVM Kernel RBF.	59
Gambar 4.14 Confusion Matrix Training SVM Kernel RBF.	59
Gambar 4.15 Evaluasi Confusion Matrix Testing SVM Linear.	61
Gambar 4.16 Evaluasi Confusion Matrix Testing SVM RBF.	62
Gambar 4.17 Evaluasi Confusion Matrix Grid Search.	63
Gambar 4.18 WordCloud Sentimen positif.	65
Gambar 4.19 WordCloud Sentimen Negatif.	66
Gambar 4.20 Tampilan UI Analisis Sentimen Ulasan JogjaKita.	68
Gambar 4.21 Hasil Analisis Sentimen Positif.	68
Gambar 4.22 Hasil Analisis Sentimen Negatif.	69

DAFTAR LAMPIRAN

Source code web scraping.....	74
Source code Analisis sentimen menggunakan SVM	75
Link Implementasi Sistem	80



DAFTAR LAMBANG DAN SINGKATAN



AI	Artificial Intelligence
C	Parameter regulasi pada algoritma SVM
CV	Cross Validation
SVM	Support Vector Machines
FN	False Negative
FP	False Positive
γ (gamma)	Parameter kernel RBF pada SVM
IDF	Inverse Document Frequency
NLP	Natural Language Processing
RBF	Radial Basis Function
TF-IDF	Term Frequency – Inverse Document Frequency
TP	True Positive
TN	True Negative

DAFTAR ISTILAH

Analisis Sentimen	Proses mengidentifikasi dan mengklasifikasikan opini pengguna ke dalam kategori positif atau negatif berdasarkan teks ulasan.
Confusion Matrix	Tabel evaluasi yang menunjukkan jumlah prediksi benar dan salah dari suatu model klasifikasi.
Cross Validation	Teknik validasi model dengan membagi data latih ke dalam beberapa lipatan untuk memperoleh performa yang lebih stabil.
Data Latih (Training Data)	Data yang digunakan untuk melatih model klasifikasi agar dapat mempelajari pola dari data.
Data Uji (Testing Data)	Data yang digunakan untuk mengevaluasi performa model setelah proses pelatihan selesai.
Grid Search	Metode pencarian sistematis untuk menentukan kombinasi hyperparameter terbaik dari nilai yang telah ditentukan.
Hyperparameter	Parameter yang ditentukan sebelum proses pelatihan model dan tidak dipelajari langsung dari data.
Kernel	Fungsi pada SVM yang digunakan untuk memetakan data ke ruang berdimensi lebih tinggi.
Preprocessing Teks	Tahapan pembersihan dan normalisasi teks sebelum dilakukan ekstraksi fitur.
Stratified Split	Teknik pembagian data yang mempertahankan proporsi kelas pada data latih dan data uji.
Support Vector Machine	Algoritma machine learning yang digunakan untuk klasifikasi dengan mencari hyperplane pemisah terbaik.
WordCloud	Visualisasi kata berdasarkan frekuensi kemunculan untuk menggambarkan karakteristik teks.

INTISARI

Penelitian ini bertujuan untuk menganalisis dan mengevaluasi performa algoritma *Support Vector Machine* (SVM) dalam mengklasifikasikan sentimen ulasan pengguna aplikasi JogjaKita yang diperoleh dari platform Google Play Store. Data penelitian berjumlah 1.965 ulasan berbahasa Indonesia yang dikumpulkan melalui teknik *web scraping* dan dikategorikan menjadi dua kelas sentimen, yaitu positif dan negatif. Proses analisis dilakukan melalui beberapa tahapan *text preprocessing* berbasis *Natural Language Processing* (NLP) yang meliputi *case folding*, *cleaning*, *tokenizing*, *stopword removal*, dan *stemming* untuk menghasilkan teks yang bersih dan seragam. Selanjutnya, data teks direpresentasikan ke dalam bentuk numerik menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF). Model SVM diuji dengan dua jenis kernel, yaitu Linear dan *Radial Basis Function* (RBF), serta dilakukan penalaan parameter menggunakan metode *Grid Search* untuk memperoleh hasil optimal. Evaluasi kinerja model menggunakan metrik akurasi, presisi, *recall*, dan *f1-score*. Hasil pengujian menunjukkan bahwa algoritma SVM dengan kernel RBF dan optimasi parameter *Grid Search* tanpa penyeimbangan data memberikan performa terbaik dengan tingkat akurasi sebesar 87,02%, *precision* rata-rata 0,86, *recall* 0,87, dan *f1-score* 0,87. Temuan ini membuktikan bahwa metode SVM efektif diterapkan untuk analisis sentimen teks berbahasa Indonesia dan dapat menjadi acuan dalam pengembangan model analisis opini publik pada aplikasi layanan publik digital.

Kata kunci: analisis sentimen, machine learning, SVM, Grid Search, TF-IDF, NLP, JogjaKita.

ABSTRACT

This research aims to analyze and evaluate the performance of the Support Vector Machine (SVM) algorithm in classifying user review sentiments of the JogjaKita application collected from the Google Play Store platform. The dataset comprises 1,965 Indonesian-language reviews obtained through a web scraping technique and categorized into two sentiment classes: positive and negative. The analytical process involved a series of text preprocessing procedures based on Natural Language Processing (NLP), including case folding, cleaning, tokenizing, stopword removal, and stemming, to ensure standardized and noise-free textual data. Subsequently, the text data were transformed into numerical representations using the Term Frequency-Inverse Document Frequency (TF-IDF) method. The SVM model was evaluated using two kernel types, namely Linear and Radial Basis Function (RBF), with parameter optimization performed through the Grid Search method to obtain optimal classification performance. Model evaluation employed several metrics, including accuracy, precision, recall, and f1-score. The experimental results revealed that the SVM algorithm with the RBF kernel and parameter optimization via Grid Search without data balancing achieved the best performance, with an accuracy of 87.02%, an average precision of 0.86, a recall of 0.87, and an f1-score of 0.87. These findings demonstrate that the SVM method, optimized through Grid Search, is effective for sentiment analysis of Indonesian textual data and can serve as a methodological reference for developing sentiment classification models in public digital service applications.

Keyword: sentiment analysis, machine learning, SVM, Grid Search, TF-IDF, NLP, JogjaKita