

**IMPLEMENTASI ALGORITMA SVM UNTUK
MEMPREDIKSI KEPOPULERAN LAGU BERDASARKAN
ANALISIS SENTIMEN LIRIK**

LAPORAN NON-REGULER

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



Disusun oleh :

QUIIN LATIFAH ALMATIN LUBIS

21.11.4069

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

**IMPLEMENTASI ALGORITMA SVM UNTUK MEMPREDIKSI
KEPOPULERAN LAGU BERDASARKAN ANALISIS SENTIMEN
LIRIK**

LAPORAN NON-REGULER

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



Disusun oleh :

QUIIN LATIFAH ALMATIN LUBIS

21.11.4069

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

HALAMAN PERSETUJUAN

JALUR NON-REGULER (SCIENTIST)

IMPLEMENTASI ALGORITMA SVM UNTUK MEMPREDIKSI
KEPOPULERAN LAGU BERDASARKAN ANALISIS SENTIMEN LIRIK

yang disusun dan diajukan oleh

Quinn Latifah Almatin Lubis

21.11.4069

telah disetujui oleh Dosen Pembimbing
pada tanggal 19 Februari 2025

Dosen Pembimbing,



Arif Akbarul Huda, S.Si., M.Eng.

NIK. 190302287

HALAMAN PENGESAHAN

JALUR NON-REGULER (SCIENTIST)

IMPLEMENTASI ALGORITMA SVM UNTUK MEMPREDIKSI
KEPOPULERAN LAGU BERDASARKAN ANALISIS SENTIMEN LIRIK

yang disusun dan diajukan oleh

Quinn Latifah Almatin Lubis

21.11.4069

Telah dipertahankan di depan Dewan Penguji
pada tanggal 19 Februari 2025

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Theophilus Bayu Sasongko, S.Kom., M.Eng.
NIK. 190302375

Majid Rahardil, S.Kom., M.Eng.
NIK. 190302393

Arif Akbarul Huda, S.Si., M.Eng.
NIK. 190302287

Laporan ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 19 Februari 2025

DEKAN FAKULTAS ILMU KOMPUTER



Hanif Al Fatta, S.Kom., M.Kom., Ph.D.
NIK. 190302096

HALAMAN PERNYATAAN KEASLIAN KARYA

Yang bertandatangan di bawah ini,

Nama mahasiswa : Quiin Latifah Almatin Lubis

NIM : 21.11.4069

Menyatakan bahwa laporan dengan judul berikut:

Implementasi Algoritma SVM untuk Memprediksi Kepopuleran Lagu berdasarkan Analisis Sentimen Lirik

Dosen Pembimbing : Arif Akbarul Huda, S.Si., M.Eng.

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 kegiatan 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 ketidak-benaran 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, 19 Februari 2025

Yang Menyatakan,



Quiin Latifah Almatin Lubis

HALAMAN PERSEMBAHAN

Dengan rasa hormat dan syukur, skripsi ini saya persembahkan kepada kedua orang tua tercinta yang senantiasa memberikan doa, kasih sayang serta dukungan tanpa henti; kakak-kakak saya yang selalu memberikan semangat dan dukungan moral maupun material; seluruh teman-teman yang telah memberikan dorongan, kebersamaan dan kenangan berharga selama perjalanan akademik ini; serta almamater saya Universitas Amikom Yogyakarta yang telah menjadi tempat menimba ilmu dan mengembangkan diri.



KATA PENGANTAR

Puji syukur kehadiran Allah Swt. atas rahmat dan karunia-Nya, sehingga saya dapat menyelesaikan skripsi ini sebagai salah satu syarat untuk memperoleh gelar Sarjana di Universitas Amikom Yogyakarta.

Dalam penyusunan skripsi ini, saya mendapatkan dukungan, bimbingan, serta bantuan dari berbagai pihak. Oleh karena itu, dengan penuh rasa hormat, saya ingin mengucapkan terima kasih kepada:

1. Bapak Arif Akbarul Huda, S.Si., M.Eng., selaku Dosen Pembimbing yang dengan penuh kesabaran telah memberikan bimbingan, masukan, serta motivasi selama proses penyusunan skripsi ini.
2. Bapak Theophilus Bayu Sasongko, S.Kom., M.Eng. dan Bapak Majid Rahardi, S.Kom., M.Eng., selaku Dosen Penguji yang telah memberikan kritik dan saran yang sangat berarti dalam penyempurnaan skripsi ini.
3. Kedua orang tua, yang senantiasa memberikan doa, dukungan, dan motivasi dalam setiap langkah yang saya tempuh.
4. Seluruh pihak yang telah membantu, baik secara langsung maupun tidak langsung dalam penyusunan skripsi ini.

Saya menyadari bahwa skripsi ini masih jauh dari sempurna. Oleh karena itu, saya terbuka terhadap kritik dan saran yang membangun demi perbaikan di masa mendatang. Semoga skripsi ini dapat memberikan manfaat bagi semua pihak yang membacanya.

Yogyakarta, 19 Februari 2025

Penulis

DAFTAR ISI

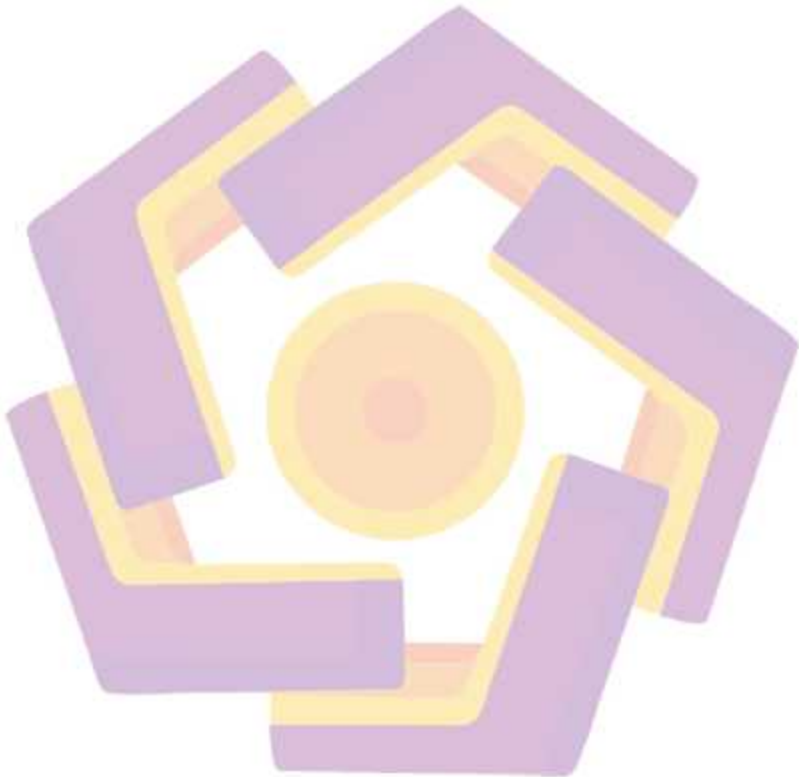
HALAMAN JUDUL.....	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN.....	iii
HALAMAN PERNYATAAN KEASLIAN KARYA.....	iv
HALAMAN PERSEMBAHAN.....	v
KATA PENGANTAR.....	vi
DAFTAR ISI.....	vii
DAFTAR TABEL.....	ix
DAFTAR GAMBAR.....	x
DAFTAR LAMPIRAN.....	xi
INTISARI.....	xii
ABSTRACT.....	xiii
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang.....	1
BAB II TINJAUAN PUSTAKA.....	3
2.1. Studi Literatur.....	3
2.2. Landasan Teori.....	4
BAB III METODE PENELITIAN.....	7
3.1. <i>Scraping</i> Data.....	7
3.2. <i>Labeling</i> Data.....	7
3.3. <i>Pre-processing</i> Data.....	8
3.4. Ekstraksi Fitur (TF-IDF).....	8
3.5. <i>Handling Imbalanced Data</i> (SMOTE).....	9
3.6. <i>Support Vector Machine</i>	10
3.7. Evaluasi Model.....	10
3.8. Visualisasi.....	11
3.9. Implementasi Model.....	11
BAB IV HASIL DAN PEMBAHASAN.....	12
4.1. Hasil <i>Scraping</i> Data.....	12
4.2. Hasil <i>Labeling</i> Data.....	12
4.3. Hasil <i>Pre-processing</i> Data.....	13
4.4. Hasil TF-IDF.....	14
4.5. Hasil SMOTE.....	15
4.6. <i>Support Vector Machine</i>	16
4.7. Evaluasi Model.....	17

4.8. Visualisasi.....	18
4.9. Implementasi Model.....	20
BAB V PENUTUP.....	21
5.1. Kesimpulan.....	21
5.2. Saran.....	21
REFERENSI.....	22
CURRICULUM VITAE.....	24
LAMPIRAN DAN BUKTI PENDUKUNG.....	25
a. Identitas Jurnal.....	25
b. Surat Rekomendasi Non-Regular.....	25
c. <i>Letter of Acceptance</i> (LOA).....	26
d. Lembar <i>Review</i>	27
e. Bukti Terbit/Terindex.....	29
f. Bukti pembayaran.....	29



DAFTAR TABEL

Tabel 4.1. Hasil <i>Pre-processing</i>	14
Tabel 4.2. Perbandingan Hasil Sebelum dan Setelah <i>Hyperparameter Tuning</i>	18

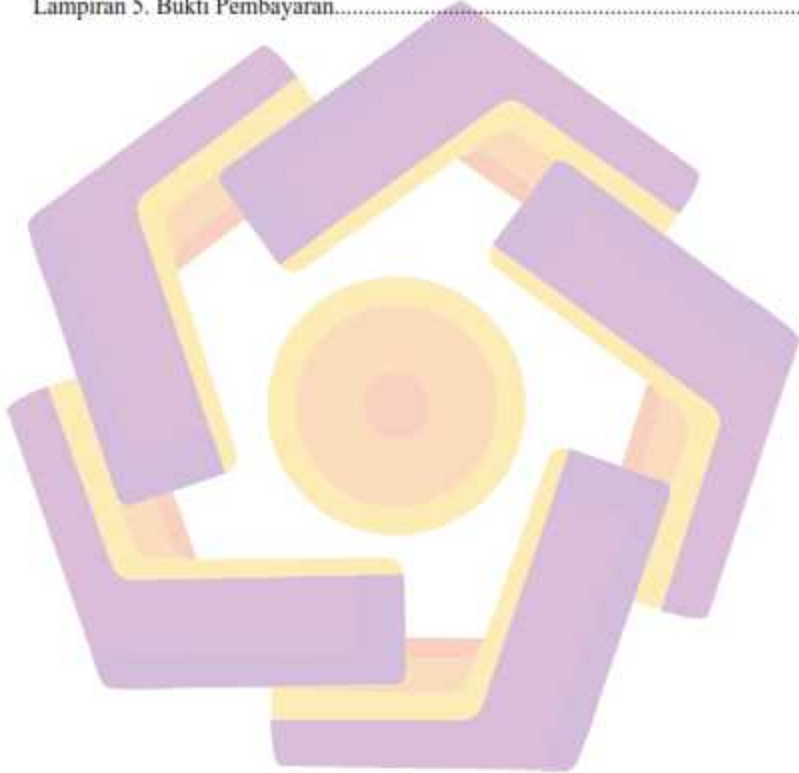


DAFTAR GAMBAR

Gambar 3.1. Alur Penelitian.....	7
Gambar 3.2. Algoritma SVM.....	10
Gambar 4.1. Hasil <i>Scraping</i>	12
Gambar 4.2. Statistik Deskriptif.....	12
Gambar 4.3. Hasil <i>Labeling</i>	13
Gambar 4.4. Hasil TF-IDF.....	15
Gambar 4.5. Distribusi Data Sebelum SMOTE.....	16
Gambar 4.6. Distribusi Data Setelah SMOTE.....	16
Gambar 4.7. Grafik Hasil <i>GridSearchCV</i>	17
Gambar 4.8. <i>Confusion Matrix</i>	18
Gambar 4.9. <i>Word Cloud</i> Lirik Sangat Populer.....	19
Gambar 4.10. <i>Word Cloud</i> Lirik Lumayan Populer.....	19
Gambar 4.11. <i>Word Cloud</i> Lirik Kurang Populer.....	19
Gambar 4.12. Hasil Implementasi Model.....	20

DAFTAR LAMPIRAN

Lampiran 1. Surat Rekomendasi Non-Reguler.....	25
Lampiran 2. <i>Letter of Acceptance</i> (LOA).....	26
Lampiran 3. Lembar <i>Review</i>	28
Lampiran 4. Bukti Terbit.....	29
Lampiran 5. Bukti Pembayaran.....	29



INTISARI

Musisi independen menghadapi tantangan besar dalam meningkatkan visibilitas dan daya tarik karya mereka di tengah persaingan ketat pada platform streaming musik. Meskipun berbagai penelitian telah dilakukan untuk menganalisis dan memprediksi kepopuleran lagu, sebagian besar penelitian tersebut berfokus pada lagu berbahasa Inggris. Hal ini menciptakan kesenjangan penelitian terhadap lagu berbahasa Indonesia, terutama dalam konteks prediksi popularitas berdasarkan lirik. Dataset yang digunakan mencakup 652 lagu berbahasa Indonesia dari tahun 2017 hingga 2024. Metodologi penelitian meliputi *pre-processing* data, ekstraksi fitur menggunakan TF-IDF, penanganan ketidakseimbangan data dengan SMOTE, implementasi SVM, dan optimasi model. Hasil penelitian menunjukkan peningkatan akurasi model dari 84% menjadi 89% setelah optimasi parameter menggunakan *GridSearchCV*. Dalam evaluasi model dengan *cross-validation* sebanyak 5-fold, diperoleh rata-rata akurasi 86,19% dengan standar deviasi 0,90%. Metrik *precision*, *recall*, dan *F1-score* untuk kelas Kurang Populer adalah 0,98, 0,85, dan 0,91; untuk kelas Lumayan Populer 0,79, 0,95, dan 0,86; dan untuk kelas Sangat Populer 0,92, 0,86, dan 0,89. Implementasi model dalam aplikasi Streamlit memungkinkan prediksi popularitas lagu berdasarkan lirik, memberikan wawasan berharga bagi musisi dalam memilih diksi yang berpotensi meningkatkan popularitas lagu mereka.

Kata kunci: Analisis Sentimen, Lirik Lagu, *Support Vector Machine*, Prediksi Popularitas.

ABSTRACT

Independent musicians face significant challenges in enhancing the visibility and appeal of their work amid intense competition on music streaming platforms. Although numerous studies have been conducted to analyze and predict song popularity, most of them focus on English-language songs. This creates a research gap for Indonesian-language songs, particularly in the context of predicting popularity based on lyrics. The dataset used includes 652 Indonesian songs from 2017 to 2024. The research methodology includes data pre-processing, feature extraction using TF-IDF, handling data imbalance with SMOTE, implementing SVM, and model optimization. The results show an improvement in model accuracy from 84% to 89% after parameter optimization using GridSearchCV. In the model evaluation with 5-fold cross-validation, an average accuracy of 86.19% with a standard deviation of 0.90% was obtained. Precision, Recall, and F1-score metrics for the Less Popular class are 0.98, 0.85, and 0.91; for the Moderately Popular class, 0.79, 0.95, and 0.86; and for the Very Popular class, 0.92, 0.86, and 0.89. The implementation of the model in a Streamlit application allows for the prediction of song popularity based on lyrics, providing valuable insights for musicians in choosing word choices that can potentially increase the popularity of their songs.

Keyword: Sentiment Analysis, Song Lyrics, Support Vector Machine, Popularity Prediction.