

**ANALISIS SENTIMEN ULASAN PENGGUNA JAKLINGKO PADA YOUTUBE
MENGUNAKAN ALGORITMA SVM DAN NAÏVE BAYES**

SKRIPSI SARJANA SISTEM INFORMASI



Oleh :

Isnando Noverdy Catur Ahiruwa

217006416051

FAKULTAS TEKNOLOGI KOMUNIKASI DAN INFORMATIKA

PROGRAM STUDI SISTEM INFORMASI

UNIVERSITAS NASIONAL

2025

**ANALISIS SENTIMEN ULASAN PENGGUNA JAKLINGKO PADA YOUTUBE
MENGUNAKAN ALGORITMA SVM DAN NAÏVE BAYES**

SKRIPSI

Karya ilmiah sebagai salah satu syarat untuk memperoleh gelar Sarjana Sistem
Informasi dari Fakultas Teknologi Komunikasi dan Informatika



Oleh :

Isnando Noverdy Catur Ahiruwa

217006416051

FAKULTAS TEKNOLOGI KOMUNIKASI DAN INFORMATIKA

PROGRAM STUDI SISTEM INFORMASI

UNIVERSITAS NASIONAL

2025

HALAMAN PENGESAHAN

TUGAS AKHIR

ANALISIS SENTIMEN ULASAN PENGGUNA JAKLINGKO PADA
YOUTUBE MENGGUNAKAN ALGORITMA SVM DAN NAÏVE

BAYES



Isnando Noverdy Catur Ahiruwa
217006416051

Dosen Pembimbing I

(Dr. Arie Gunawan, S.Kom., MMSI.)

PERNYATAAN KEASLIAN TUGAS AKHIR

Saya menyatakan dengan sesungguhnya bahwa Tugas Akhir dengan judul :

ANALISIS SENTIMEN ULASAN PENGGUNA JAKLINGKO PADA YOUTUBE MENGGUNAKAN ALGORITMA SVM DAN NAÏVE BAYES

Yang dibuat untuk melengkapi salah satu persyaratan menjadi Sarjana Komputer pada Program Studi Sistem Informasi, Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional, sebagaimana yang saya ketahui adalah bukan merupakan tiruan atau publikasi dari Tugas Akhir yang pernah diajukan atau dipakai untuk mendapatkan gelar di lingkungan Universitas Nasional maupun perguruan tinggi atau instansi lainnya, kecuali pada bagian – bagian tertentu yang menjadi sumber Informasi atau acuan yang dicantumkan sebagaimana mestinya.

Jakarta, 28 febuari 2025



Isnando Noverdy Catur Ahinuwa

217006416051

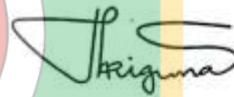
LEMBAR PERSETUJUAN REVIEW AKHIR

Tugas Akhir dengan judul :

ANALISIS SENTIMEN ULASAN PENGGUNA JAKLINGKO PADA YOUTUBE MENGGUNAKAN ALGORITMA SVM DAN NAÏVE BAYES

Dibuat untuk melengkapi salah satu persyaratan menjadi Sarjana Komputer pada Program Studi Sistem Informasi, Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional. Tugas Akhir ini diujikan pada Sidang Akhir Semester Ganjil 2024-2025 pada tanggal 26 Februari 2025

Dosen Pembimbing 1



Dr. Arie Gunawan, S.Kom., MMSI.

NIDN. 0410047808

Ketua Program Studi



Dr. Andrianiingsih, S. Kom., MMSI.

NIDN. 0303097902

LEMBAR PERSETUJUAN JUDUL YANG TIDAK ATAU YANG DIREVISI

Nama : Isnando Noverdy Catur Ahiruwa
 NPM : 217006416051
 Fakultas/Akademi : Fakultas Teknologi Komunikasi dan Informatika
 Program Studi : Sistem Informasi
 Tanggal Sidang : 26 Februari 2025

JUDUL DALAM BAHASA INDONESIA :

ANALISIS SENTIMEN ULASAN PENGGUNA JAKLINGKO PADA
 YOUTUBE MENGGUNAKAN ALGORITMA SVM DAN NAÏVE
 BAYES

JUDUL DALAM BAHASA INGGRIS :

SENTIMENT ANALYSIS OF JAKLINGKO USER REVIEWS ON
 YOUTUBE USING SVM AND NAÏVE BAYES ALGORITHMS

TANDA TANGAN DAN TANGGAL		
Pembimbing I	Ka. Prodi	Mahasiswa
TGL : 28-02-2025	TGL : 28-02-2025	TGL : 28-02-2025
	 Dr. Andriyungsih, MMCI	

ABSTRAK

Transportasi publik yang terintegrasi menjadi kebutuhan utama dalam menunjang mobilitas masyarakat di Jakarta. JakLingko hadir sebagai solusi yang menghubungkan berbagai moda transportasi dalam satu sistem pembayaran yang efisien. Penelitian ini bertujuan untuk menganalisis sentimen ulasan pengguna terhadap layanan JakLingko di platform YouTube menggunakan algoritma Support Vector Machine (SVM) dan Naïve Bayes. Metode yang digunakan meliputi pengumpulan data melalui web scraping, preprocessing teks, serta klasifikasi sentimen dengan kedua algoritma tersebut.

Menunjukkan bahwa sentimen pengguna terhadap JakLingko cenderung positif, dengan mayoritas ulasan memberikan apresiasi terhadap integrasi layanan. Namun, terdapat juga sentimen negatif yang menyoroti kendala teknis dalam penggunaan aplikasi dan layanan operasional. Dari segi performa, algoritma SVM menunjukkan akurasi yang lebih tinggi dibandingkan Naïve Bayes, dengan nilai 96% berbanding 85%. Hal ini menunjukkan bahwa SVM lebih andal dalam mengklasifikasikan sentimen berdasarkan pola data yang kompleks.

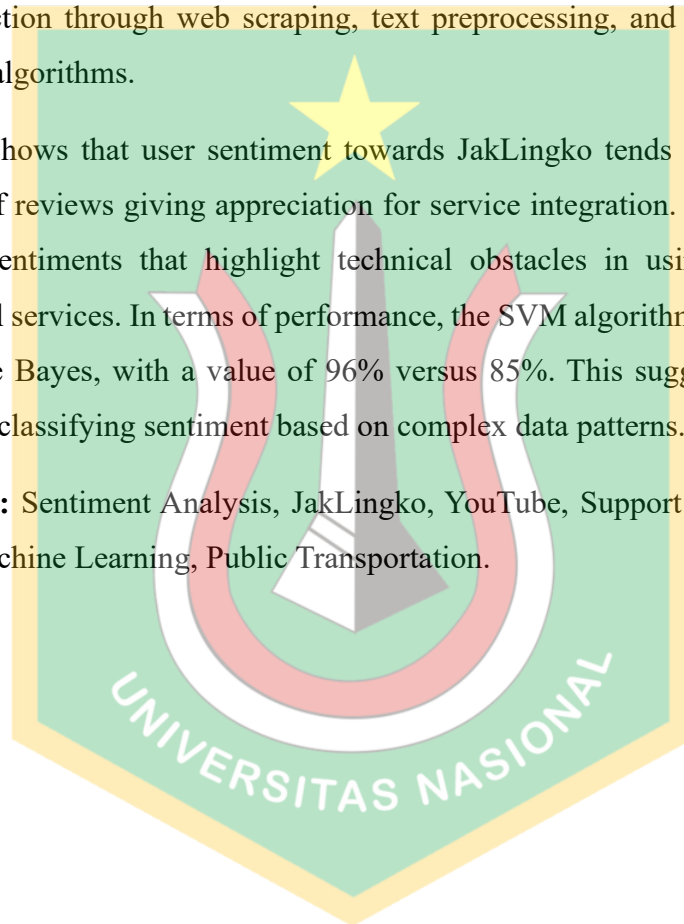
Kata kunci: Analisis Sentimen, JakLingko, YouTube, Support Vector Machine, Naïve Bayes, Machine Learning, Transportasi Publik.

ABSTRACT

Integrated public transportation is a major requirement in supporting people's mobility in Jakarta. JakLingko comes as a solution that connects various modes of transportation in one efficient payment system. This research aims to analyze the sentiment of user reviews of JakLingko services on the YouTube platform using the Support Vector Machine (SVM) and Naïve Bayes algorithms. The methods used include data collection through web scraping, text preprocessing, and sentiment classification with both algorithms.

It shows that user sentiment towards JakLingko tends to be positive, with the majority of reviews giving appreciation for service integration. However, there are also negative sentiments that highlight technical obstacles in using the application and operational services. In terms of performance, the SVM algorithm shows higher accuracy than Naïve Bayes, with a value of 96% versus 85%. This suggests that SVM is more reliable in classifying sentiment based on complex data patterns.

Keywords: Sentiment Analysis, JakLingko, YouTube, Support Vector Machine, Naïve Bayes, Machine Learning, Public Transportation.



KATA PENGANTAR

Puji syukur ke hadirat Tuhan Yang Maha Esa atas segala rahmat dan karunia-Nya sehingga skripsi ini dapat diselesaikan dengan baik. Skripsi yang berjudul “ANALISIS SENTIMEN ULASAN PENGGUNA JAKLINGKO PADA YOUTUBE MENGGUNAKAN ALGORITMA SVM DAN NAÏVE BAYES” disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana di Fakultas Teknologi Komunikasi dan Informatika, Program Studi Sistem Informasi, Universitas Nasional.

Penyusunan skripsi ini tidak terlepas dari bantuan, bimbingan, serta dukungan dari berbagai pihak. Oleh karena itu, dengan segala hormat dan rasa terima kasih yang mendalam, penulis menyampaikan apresiasi kepada:

1. Bapak Dr. El Amry Bermawi Putera, M.A., selaku Rektor Universitas Nasional yang telah memberikan fasilitas dan kesempatan bagi mahasiswa dalam menempuh pendidikan tinggi dengan baik.
2. Bapak Dr. Agung Triayudi, S.Kom., M.Kom, selaku Dekan Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional yang telah memberikan kesempatan dan fasilitas dalam penyelesaian studi.
3. Ibu Dr. Andrianingsih, S.Kom.,MMSI., selaku Ketua Program Studi Sistem Informasi Universitas Nasional atas arahan dan kebijakan yang diberikan selama masa perkuliahan.
4. Ibu Dr. Arie Gunawan, S.Kom.,MMSI., selaku dosen pembimbing yang telah memberikan bimbingan, masukan, serta motivasi yang sangat berharga dalam proses penyusunan skripsi ini.
5. Bapak/Ibu Dosen Universitas Nasional, yang telah memberikan ilmu, wawasan, dan pengalaman berharga selama masa perkuliahan.
6. Keluarga tercinta ibu, ayah, kakak, adik, atas segala doa, dukungan, serta pengorbanan yang tak ternilai harganya dalam mendukung penyelesaian studi ini.

7. Teman-teman seperjuangan Tio, Fauzan, Irzan, Aji, Rendi, Putra, Rival, Lintar, Dika, I Gede, Gerald, Adam, Daffa, Aqshal yang selalu memberikan semangat, motivasi, serta dukungan moral dalam menjalani proses akademik dan penelitian ini.
8. Seluruh pihak yang tidak dapat disebutkan satu per satu, yang telah membantu secara langsung maupun tidak langsung dalam penyelesaian skripsi ini.

Penulis menyadari bahwa karya ini dapat memberikan manfaat bagi perkembangan ilmu pengetahuan, khususnya dalam bidang analisis sentimen dan pemanfaatan algoritma pembelajaran mesin dalam dunia digital .skripsi ini masih memiliki keterbatasan dan kekurangan. Oleh karena itu, segala kritik dan saran yang membangun sangat diharapkan guna penyempurnaan penelitian ini.



Jakarta, 26 Februari 2025

Peneliti

Isnando Noverdy Catur Ahiurwa

DAFTAR ISI

BAB I PENDAHULUAN.....	17
1.1 Latar Belakang.....	17
1.2 Rumusan Masalah.....	21
1.3 Tujuan Penelitian.....	21
1.4 Batasan Masalah.....	21
1.5 Kontribusi Peneliti.....	22
BAB II TINJAUAN PUSTAKA.....	23
2.1 Studi Literatur.....	23
2.2 Landasan Teori.....	30
2.2.1 Analisis Sentimen.....	30
2.2.2 Support Vector Machine (SVM).....	31
2.2.3 Naïve Bayes.....	31
2.2.4 Jaklingko.....	32
2.3 Preprocessing Data.....	33
2.3.1 Case Folding.....	33
2.3.2 Cleansing.....	33
2.3.3 Normalization.....	34
2.3.4 Tokenisasi.....	34
2.3.5 Filtering (Stopwords Removal).....	34
2.3.6 Stemming.....	34
2.4 Pelabelan Sentimen.....	34
2.5 Pembobotan Kata.....	35
2.5.1 Representasi Data.....	35

2.5.2	Pembagian Dataset.....	35
2.6	Evaluasi Model	36
2.7	Visualisasi Data	36
2.8	Alat dan Teknologi Pendukung	36
BAB III METODE PENELITIAN		39
3.1	Lokasi Penelitian.....	39
3.2	Waktu Penelitian.....	39
3.3	Objek Penelitian.....	40
3.4	Fokus Penelitian.....	41
3.5	Sumber Data.....	41
3.6	Penyimpanan Data	42
3.7	Teknik Pengumpulan Data.....	42
3.7.1	Scraping Data.....	42
3.8	Tahap Preprocessing Data	43
3.8.1	Case Folding	43
3.8.2	Cleansing.....	44
3.8.3	Normalization	44
3.8.4	Tokenisasi	44
3.8.5	Filtering (Stopwords Removal).....	45
3.8.6	Stemming	45
3.9	Labelling Sentimen	46
3.10	Pembobotan Kata	47
3.10.1	Term Frequency-Inverse Document Frequency	47
3.10.2	Pembagian Dataset.....	49
3.11	Implementasi Algoritma	50

3.11.1	Naïve Bayes	50
3.11.2	Support Vector Machine	51
3.12	Evaluasi Model	53
3.13	Desiain Alur Penelitian	55
BAB IV HASIL DAN PEMBAHASAN		57
4.1	Teknik Pengumpulan Data	57
4.1.1	Scraping Data	57
4.2	Preprocessing Data	57
4.2.1	Case Folding	57
4.2.2	Cleansing	58
4.2.3	Normalization	59
4.2.4	Tokenizing	60
4.2.5	Filtering (Stop Word Removal)	61
4.2.6	Stemming	62
4.3	Pelabelan InsetLexicon	63
4.4	Pembobotan Kata	64
4.4.1	Ekstraksi Fitur TF-IDF	64
4.4.2	Perhitungan Manual TF	65
4.4.3	Perhitungan Manual IDF	69
4.4.4	Perhitungan Manual TF-IDF	71
4.4.5	Pembagian Data	72
4.5	Perhitungan Manual Naïve Bayes	73
4.5.1	Total Kata dalam Setiap Kelas	75
4.5.2	Probabilitas Awal (Prior Probability)	75
4.5.3	Probabilitas Likelihood	75

4.5.4	Probabilitas Posterior	76
4.6	Perhitungan Manual Suport Vector Machine (SVM)	77
4.7	Evaluasi Model	79
4.7.1	Naïve Bayes	79
4.7.2	Suport Vector Machine (SVM).....	80
4.8	Visualisasi Data	82
4.8.1	Diagram Batang	82
4.8.2	World Cloud.....	84
BAB V KESIMPULAN DAN SARAN		88
5.1	Kesimpulan	88
5.2	Saran	88
DAFTAR PUSTAKA		90



DAFTAR GAMBAR

Gambar 2. 1 Aplokasi Jaklingko	32
Gambar 2. 2 Jaklingko	33
Gambar 3. 1 Scraping Data	42
Gambar 3. 2 Preprocessing Data	43
Gambar 3. 3 Pelabelan InsetLexicon	46
Gambar 3. 4 TF-IDF	47
Gambar 3. 5 Pembagian Dataset	49
Gambar 3. 6 Implementasi Algoritma	50
Gambar 3. 7 Evaluasi Model	53
Gambar 3. 8 Confusion Matrix dalam Analisis Sentimen	53
Gambar 3. 9 Desain Alur Penelitian	55
Gambar 4. 1 Scraping (Pengambilan Data)	57
Gambar 4. 2 Pembagian Data	72
Gambar 4. 3 Confusion Matrix klasifikasi Naive Bayes	79
Gambar 4. 4 Confusion Matrix klasifikasi SVM	80
Gambar 4. 5 Diagram Batang	82
Gambar 4. 6 Top 10 Word Positif	83
Gambar 4. 7 Top 10 Word Negatif	84
Gambar 4. 8 Top 10 Word Netral	84
Gambar 4. 9 World Cloud Positif	85
Gambar 4. 10 World Cloud Negatif	85
Gambar 4. 11 World Cloud Netral	86
Gambar 4. 12 Streamlit	87

DAFTAR TABEL

Tabel 2. 1 Penelitian Terdahulu.....	23
Tabel 3. 1 Waktu Penelitian.....	39
Tabel 4. 1 Case Folding.....	57
Tabel 4. 2 Normalization.....	59
Tabel 4. 3 Tokenizing	60
Tabel 4. 4 Filtering	61
Tabel 4. 5 Stemming	62
Tabel 4. 6 Pelabelan InsetLexicon	63
Tabel 4. 7 Ekstraksi TF-IDF.....	64
Tabel 4. 8 Menentukan Nilai tf (sudah dilakukan proses stemming).....	65
Tabel 4. 9 Perhitungan Manuak Term Frequency	66
Tabel 4. 10 Menentukan Nilai IDF	69
Tabel 4. 11 Menentukan TF-IDF.....	71
Tabel 4. 12 Menentukan Term.....	73
Tabel 4. 13 Pembobotan Kata	73
Tabel 4. 14 Perhitungan Likelihood.....	75
Tabel 4. 15 Representasi Teks dengan TF-IDF	77
Tabel 4. 16 Hyperplane	77