

**ANALISIS SENTIMEN TWEETS DUKUNGAN PALESTINA PADA  
MEDIA SOSIAL X MENGGUNAKAN ALGORITMA SUPPORT  
VECTOR MACHINE DAN NAIVE BAYES**

**SKRIPSI**

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



disusun oleh

**MUHAMMAD REZZHA RIAMRIZAL ARSYA**

**22.21.1545**

Kepada

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

**2024**

**ANALISIS SENTIMEN TWEETS DUKUNGAN PALESTINA PADA  
MEDIA SOSIAL X MENGGUNAKAN ALGORITMA SUPPORT  
VECTOR MACHINE DAN NAIVE BAYES**

**SKRIPSI**

untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi Informatika



disusun oleh

**MUHAMMAD REZZHA RIAMRIZAL ARSYA**  
**22.21.1545**

Kepada

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

**2024**

HALAMAN PERSETUJUAN

SKRIPSI

**Analisis Sentimen Tweets Dukungan Palestina pada Media Sosial X  
Menggunakan Algoritma Support Vector Machine dan Naive Bayes**


yang disusun dan diajukan oleh

**Muhammad Rezza Riamrizal Arsy**

22.21.1545

telah disetujui oleh Dosen Pembimbing Skripsi  
pada tanggal 20 Februari 2024

Dosen Pembimbing,



Majid Rahardi, S.Kom., M.Eng  
NIK: 190302393

HALAMAN PENGESAHAN

SKRIPSI

Analisis Sentimen Tweets Dukungan Palestina pada Media Sosial X  
Menggunakan Algoritma Support Vector Machine dan Naive Bayes

yang disusun dan diajukan oleh

**Muhammad Rezha Riameizal Arsyah**

22.21.1545

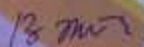
Telah dipertahankan di depan Dewan Penguji  
pada tanggal 20 Februari 2024.

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Barika Satya, M.Kom  
NIK. 190302126



Ali Mustopa, M.Kom  
NIK. 190302192



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

Skrripsi ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Sarjana Komputer  
Tanggal 20 Februari 2024.

DEKAN FAKULTAS ILMU KOMPUTER



Hanif Al Fatta, S.Kom., M.Kom., PhD.  
NIK. 190302096

## HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Muhammad Rezza Riamrizal Arsyah  
NIM : 22.21.1545

Menyatakan bahwa Skripsi dengan judul berikut:

**Analisis Sentimen Tweets Dukungan Palestina pada Media Sosial X  
Menggunakan Algoritma Support Vector Machine dan Naive Bayes**

Dosen Pembimbing : Majid Rahardi, S.Kom., 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 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, 20 Februari 2024

Yang Menyatakan,



Muhammad Rezza Riamrizal Arsyah

## HALAMAN PERSEMBAHAN

Terima kasih yang tak terhingga kepada keluarga-keluarga saya, khususnya Ibu dan Nenek, yang membesarkan, memberikan dukungan moral dan motivasi dalam setiap langkah perjalanan hidup. Ayah yang telah membiayai dan menyekolahkan, menjadi tiang yang kokoh dan penopang utama untuk meraih impian. Doa dan cinta yang tulus dari keluarga adalah penyemangat terbesar dalam menyelesaikan perjalanan panjang skripsi ini. Kebersamaan dan kasih sayang yang diberikan memberikan kekuatan yang tak tergantikan, menjadi pilar yang kokoh dalam menghadapi setiap tantangan.

Ucapan terima kasih sebesar-besarnya disampaikan kepada Bapak Majid Rahardi, S.Kom., M.Eng. selaku dosen pembimbing yang sabar dan penuh dedikasi. Bapak Majid Rahardi tidak hanya seorang pembimbing, tetapi juga seorang mentor yang memberikan arahan, saran, dan wawasan yang sangat berharga. Tanpa bantuan dan bimbingan beliau, skripsi ini tidak akan mencapai bentuknya yang sekarang.



## KATA PENGANTAR

Dengan penuh rasa syukur kepada Allah SWT atas rahmat dan karunianya, ini dengan judul **“ANALISIS SENTIMEN TWEETS DUKUNGAN PALESTINA PADA MEDIA SOSIAL X MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE DAN NAIVE BAYES”**. Skripsi ini merupakan syarat kelulusan untuk mendapatkan gelar Sarjana Informatika di Universitas Amikom Yogyakarta.

1. Bapak Prof. Dr. M. Suyanto, M.M selaku Rektor Universitas Amikom Yogyakarta.
2. Bapak Majid Rahardi, S.Kom., M.Eng selaku dosen pembimbing atas bimbingannya selama ini.
3. Keluarga yang telah mendoakan dan memberi semangat dan motivasi untuk menyelesaikan skripsi ini.
4. Teman-teman seperjuangan yang tidak dapat saya sebutkan satu persatu atas dukungan, ide, dan inspirasi untuk menyelesaikan skripsi ini.
5. Universitas Amikom Yogyakarta, atas fasilitas dan pengalaman belajar yang telah membentuk penulis menjadi pribadi yang lebih berkembang.

Akhir kata, penulis menyampaikan apresiasi yang tinggi kepada semua pihak yang telah turut berperan dalam keberhasilan penulisan skripsi ini. Semoga hasil karya ini dapat memberikan manfaat dan inspirasi.

Yogyakarta, 31 Januari 2024



Penulis

## DAFTAR ISI

HALAMAN JUDUL .....	i
.....	ii
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
Latar Belakang .....	1
Rumusan Masalah .....	2
Batasan Masalah .....	2
Tujuan Penelitian .....	2
Manfaat Penelitian .....	2
Sistematika Penulisan .....	3
DAFTAR PUSTAKA .....	4



BAB II TINJAUAN PUSTAKA .....	5
2.1 Machine Learning .....	5
2.2 Analisis Sentimen .....	5
2.3 X.....	5
2.4 Konflik Palestine-Israel .....	6
2.5 Support Vector Machine .....	6
2.6 Naïve Bayes .....	7
BAB III METODE PENELITIAN .....	17
3.1 Objek Penelitian.....	17
3.2 Alur Penelitian .....	17
3.3 Metode <i>SEMMA</i> .....	18
3.4 Studi Literatur .....	20
3.5 Alat dan Bahan.....	20
BAB IV HASIL DAN PEMBAHASAN .....	22
4.1 Pengumpulan Data .....	22
4.2 Pelabelan Data .....	23
4.3 Preprocessing Data.....	24
4.3.1 <i>Cleansing</i> (Pembersihan Data).....	25
4.3.2 <i>Tokenizing</i> (Tokenisasi).....	25
4.3.3 <i>Normalization</i> (Normalisasi).....	26
4.3.4 <i>Stopword Removal</i> (Penghapusan Stopword).....	27
4.3.5 <i>Stemming</i> (Pemangkasan kata) .....	28
4.3.6 <i>Reconstruction</i> (Detokenisasi).....	29
4.4 Wordclouds .....	30
4.5 TF-IDF .....	32
4.6 SMOTE.....	33

4.7	Pembagian Data .....	33
4.8	Klasifikasi Support Vector Machine.....	34
4.9	Klasifikasi Naïve Bayes .....	36
4.10	Evaluasi.....	38
BAB V PENUTUP .....		41
5.1	Kesimpulan .....	41
5.2	Saran .....	41
DAFTAR PUSTAKA .....		43
LAMPIRAN.....		46
	Dataset.....	46
	Kode Program Crawling Data X.....	47
	Kode Program Preprocessing Data .....	47
	Kode Program Wordcloud .....	50
	Kode Program TF-IDF.....	52
	Kode Program SMOTE.....	53
	Kode Program Pembagian Data Train & Data Test.....	53
	Kode Program Klasifikasi SVM & Naive Bayes.....	54

## DAFTAR TABEL

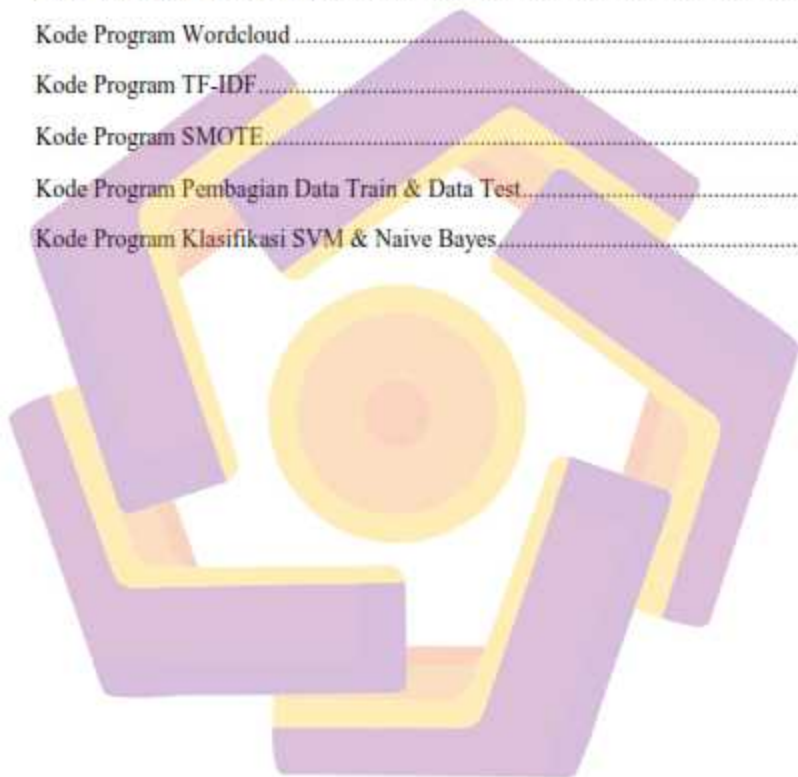
Tabel 2.1 Keaslian Penelitian .....	9
Tabel 4.1 Contoh Crawling Data .....	22
Tabel 4.2 Contoh <i>Pelabelan Data</i> .....	23
Tabel 4.3 Contoh proses <i>cleansing</i> .....	25
Tabel 4.4 Contoh proses <i>Tokenizing</i> .....	26
Tabel 4.5 Contoh kamus kata Normalisasi .....	26
Tabel 4.6 Contoh kata yang telah di Normalisasi .....	27
Tabel 4.7 Contoh kata sebelum dan sesudah <i>stopword removal</i> .....	28
Tabel 4.8 Contoh kata sebelum dan sesudah <i>Stemming</i> .....	28
Tabel 4.9 Contoh kata sebelum dan sesudah <i>Reconstruction</i> .....	29
Tabel 4.10 Jumlah kata positif.....	30
Tabel 4.11 Jumlah kata negatif.....	32
Tabel 4.12 Table Pembagian Data.....	33

## DAFTAR GAMBAR

Gambar 3.1 Flowchart Alur Penelitian .....	18
Gambar 3.2 Metode <i>SEMMA</i> .....	20
Gambar 4.1 <i>Flowchart</i> Alur Preprocessing .....	24
Gambar 4.2 Wordclouds sentimen positif .....	30
Gambar 4.3 Wordclouds sentimen negatif.....	31
Gambar 4.4 Hasil TF-IDF .....	32
Gambar 4.5 Hasil sebelum dan sesudah <i>SMOTE</i> .....	33
Gambar 4.6 Metode SVM Linier dan hasilnya .....	34
Gambar 4.7 <i>Heatmap Confusion Matrix</i> Support Vector Machine .....	35
Gambar 4.8 <i>Classification Report</i> Support Vector Machine .....	36
Gambar 4.9 Metode <i>Multinomial NB</i> dan hasilnya.....	37
Gambar 4.10 <i>Heatmap Confusion Matrix</i> Support Vector Machine .....	37
Gambar 4.11 <i>Classification Report</i> Naïve Bayes .....	38
Gambar 4.12 Kurva ROC-AUC SVM dan Naïve Bayes .....	39
Gambar 4.13 Grafik perbandingan SVM dan Naïve Bayes.....	39

## DAFTAR LAMPIRAN

Dataset.....	46
Kode Program Crawling Data X.....	47
Kode Program Preprocessing Data.....	47
Kode Program Wordcloud.....	50
Kode Program TF-IDF.....	52
Kode Program SMOTE.....	53
Kode Program Pembagian Data Train & Data Test.....	53
Kode Program Klasifikasi SVM & Naive Bayes.....	54



## DAFTAR LAMBANG DAN SINGKATAN

Acc	Accuracy (Akurasi)
AUC	Area Under the Curve
CM	Confusion Matrix (Matriks Konfusi)
FN	False Negative (Negatif Salah)
FP	False Positive (Positif Salah)
F1	F1 Score (Skor F1)
ROC	Receiver Operating Characteristic
SVM	Support Vector Machine
SMOTE	Synthetic Minority Over-sampling Technique
TF-IDF	Term Frequency-Inverse Document Frequency
TN	True Negative (Negatif Benar)
TP	True Positive (Positif Benar)



## DAFTAR ISTILAH

Bar Charts	Diagram memvisualisasikan frekuensi atau persentase
Crawling	Mengambil data dengan cara yang sistematis
Scraping	Ekstraksi informasi dari halaman web
Sentimen	Pendapat atau perasaan yang dapat diekspresikan
Prior Probability	Munculnya kelas tanpa mempertimbangkan fitur tertentu
Hyperplane	Bidang pemisahan antara kelas dalam SVM
Feature Extraction	Proses mengidentifikasi dan memilih fitur yang relevan
Training Set	Dataset yang digunakan untuk melatih model
Testing Set	Dataset untuk menguji kinerja model
Tweet-Harvest	Pengumpulan tweet dari X untuk analisis lebih lanjut
Accuracy	Mengukur model dapat memprediksi kelas dengan benar
Precision	Mengukur sejauh mana hasil positif model adalah benar
Recall	Mengukur model mendeteksi instance positif sebenarnya
Word Clouds	Representasi visual dari kata-kata dalam teks.

## INTISARI

Media sosial, sebagai platform utama interaksi di era ini, memainkan peran penting dalam menyuarakan opini dan mendukung isu-isu global. Dengan jumlah pengguna internet di Indonesia mencapai 212,9 juta dan 60,4% di antaranya aktif di media sosial, X menjadi pusat pembicaraan terkait isu-isu sensitif, termasuk konflik Israel-Palestina. Penelitian ini bertujuan untuk menganalisis sentimen terhadap tweet yang mendukung Palestina di X, membandingkan metode Support Vector Machine (SVM) dan Naïve Bayes. Analisis sentimen pada topik ini menjadi penting mengingat sorotan dunia terhadap konflik tersebut. Hasil penelitian menunjukkan bahwa model SVM memiliki performa lebih baik dibandingkan Naïve Bayes, dengan akurasi, presisi, recall, dan F1-Score mencapai 90%, 93%, 92%, dan 92% secara berturut-turut. Kurva ROC-AUC sebesar 98% menunjukkan kemampuan model SVM dalam membedakan sentimen positif dan negatif. Perbandingan dengan penelitian sebelumnya menunjukkan bahwa penelitian ini mengungguli hasil analisis sentimen dari model SVM sebelumnya, sementara Naïve Bayes lebih unggul dalam penelitian tertentu. Dengan hasil kinerja yang superior, penelitian ini memberikan kontribusi baru dalam mengenali dan mengklasifikasikan sentimen pada dataset spesifik terkait isu konflik Israel-Palestina di media sosial X.

**Kata kunci:** Analisis Sentimen, X, Support Vector Machine, Naïve Bayes, Konflik Israel-Palestina.

## ABSTRACT

*In this era, social media, as an essential interactive platform, plays a crucial role in expressing opinions and supporting global issues. With the number of internet users in Indonesia reaching 212.9 million, and 60.4% of them actively engaging on social media, X becomes a focal point for discussions on sensitive issues, including the Israel-Palestine conflict. This research aims to analyze sentiment towards tweets expressing support for Palestine on X, comparing the Support Vector Machine (SVM) and Naïve Bayes methods. Sentiment analysis on this topic is crucial given the global attention to the conflict. The results reveal that the SVM model outperforms Naïve Bayes, achieving accuracy, precision, recall, and F1-Score of 90%, 93%, 92%, and 92%, respectively. The ROC-AUC curve of 98% demonstrates the SVM model's capability to distinguish between positive and negative sentiments. Comparison with previous research indicates that this study surpasses the sentiment analysis results of previous SVM models, while Naïve Bayes excels in specific research. With superior performance, this research provides a new contribution to recognizing and classifying sentiments on a specific dataset related to the Israel-Palestine conflict on X.*

**Keyword:** Sentiment Analysis, X, Support Vector Machine, Naïve Bayes, Israel-Palestine Conflict