

**ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP CALON  
PRESIDEN INDONESIA 2024 MENGGUNAKAN METODE SUPPORT  
VECTOR MACHINE**

**SKRIPSI**

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



disusun oleh

**LABIB AQIL SIRAJ**

**19.61.0165**

Kepada

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

**2022**

**ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP CALON  
PRESIDEN INDONESIA 2024 MENGGUNAKAN METODE SUPPORT  
VECTOR MACHINE**

**SKRIPSI**

untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi Informatika



disusun oleh

**LABIB AQIL SIRAJ**

**19.01.0165**

Kepada

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA  
2022**

**HALAMAN PERSETUJUAN**

**SKRIPSI**

**ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP CALON  
PRESIDEN INDONESIA 2024 MENGGUNAKAN METODE  
SUPPORT VECTOR MACHINE**

yang disusun dan diajukan oleh

**Labib Aqil Siraj**

**19.01.0165**

telah disetujui oleh Dosen Pembimbing Skripsi  
pada tanggal 3 Januari 2023

**Dosen Pembimbing,**

**Anna Balta, S.Kom., M.Kom.**

**NIK. 190302290**

**HALAMAN PENGESAHAN**  
**SKRIPSI**  
**ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP CALON**  
**PRESIDEN INDONESIA 2024 MENGGUNAKAN METODE**  
**SUPPORT VECTOR MACHINE**

yang disusun dan diajukan oleh

**Labib Aqil Straj**

**19.61.0165**

Telah dipertahankan di depan Dewan Penguji  
pada tanggal 18 Januari 2023

**Susunan Dewan Penguji**

**Nama Penguji**

**Tanda Tangan**

**Ariflyanto Hadinegoro, S.Kom., MT.**  
**NIK. 190302289**



**Arif Dwi Laksito, M.Kom.**  
**NIK. 1903022150**

**Anna Balta, S.Kom., M.Kom.**  
**NIK. 190302290**

Skripsi ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Sarjana Komputer  
Tanggal 18 Januari 2023

**DEKAN FAKULTAS ILMU KOMPUTER**



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

#### HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Labib Aqil Siraj  
NIM : 19.61.0165

Menyatakan bahwa Skripsi dengan judul berikut:

**ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP CALON  
PRESIDEN INDONESIA 2024 MENGGUNAKAN METODE SUPPORT  
VEKTOR MACHINE**

Dosen Pembimbing : Anas Rifa'i, S.Kom, M.Kom

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMDEOM 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 AMDEOM 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 penarikan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 18 Januari 2023

Yang Menyatakan,



Labib Aqil Siraj

## HALAMAN PERSEMBAHAN

Alhamdulillah, segala puji bagi Allah SWT, Rabb semesta alam yang telah memberikan segala kenikmatan yang tidak bisa hitung jumlahnya. Shalawat serta salam semoga senantiasa ditujukan bagi Rasulullah, keluarga, para sahabat, dan siapa saja yang meneladani mereka dengan baik hingga hari kiamat.

Terimakasih kepada orang tua saya tercinta, Edi Sukisno dan Sri Wahyuni yang telah mendidik saya, mendukung saya dalam melangkah, menemani saya setiap saat, dan setiap hari mendo'akan saya hingga saya mampu untuk menyelesaikan segala kewajiban saya di bangku pendidikan. Kepada kakak dan adik saya, Nur Wahidzatun Nafisah dan Zahro Wahyu Afifah, dan kepada yang saya cintai Fabila Fatya Putri, terimakasih atas segala dukungan dan motivasinya, semoga kita kelak dikumpulkan didalam surga – Nya.

Terimakasih kepada dosen – dosen yang telah sabar dan ikhlas dalam mendidik saya hingga mampu melewati seluruh ujian mata kuliah, semoga ilmu yang diberikan bermanfaat bagi seluruh mahasiswa Informatika Universitas Amikom Yogyakarta.

## KATA PENGANTAR

*Assalamu 'alaikum Wr.Wb.*

Segala puji bagi Allah SWT, Rabb semesta alam. Shalawat dan salam semoga senantiasa ditujukan bagi Rasulullah SAW, keluarga, para sahabat, dan siapa saja yang meneladani mereka dengan baik hingga hari kiamat.

Dalam menyelesaikan skripsi ini, banyak pihak yang telah memberikan bantuan baik secara moril, nasihat, dan semangat maupun materil. Atas segala bantuan yang telah diberikan, penulis ingin menyampaikan doa dan ucapan terimakasih yang sedalam – dalamnya kepada :

1. Prof. Dr. M. Suyanto, M.M. selaku Rektor Universitas Amikom Yogyakarta.
2. Bapak Hanif Al Fatta, S.Kom, M.Kom selaku Dekan Fakultas Ilmu Komputer
3. Ibu Anna Baita, M.Kom. selaku Dosen Pembimbing yang telah meluangkan waktu untuk membimbing, mengarahkan dan memberi masukan kepada penulis dalam pengerjaan skripsi ini hingga akhir.
4. Seluruh Dosen, Laboran, dan Staff Administrasi Jurusan Informatika Universitas Amikom Yogyakarta, terima kasih atas segala ilmu dan bimbingannya.
5. Semua pihak yang tidak bisa disebutkan satu persatu yang telah membantuk menyelesaikan skripsi ini.



Berbagai kekurangan dan kesalahan mungkin pembaca temukan dalam penulisan skripsi ini, untuk itu penulis menerima segala kritik dan saran yang membangun dari pembaca sekalian. Semoga apa yang menjadi kekurangan bisa disempurnakan oleh peneliti selanjutnya dan semoga karya ini senantiasa dapat memberi manfaat.

*Wassalamualikum Wr.Wb.*

Yogyakarta, 18 Januari 2023

Penulis

Labib Aqil Siraj

19.61.0165





## DAFTAR ISI

HALAMAN JUDUL .....	i
DAFTAR ISI.....	viii
DAFTAR TABEL.....	x
DAFTAR GAMBAR.....	xi
DAFTAR LAMPIRAN.....	xii
DAFTAR LAMBANG DAN SINGKATAN.....	xiii
INTISARI.....	xiv
ABSTRACT.....	xv
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.....	3
1.6 Sistematika Penulisan.....	3
BAB II TINJAUAN PUSTAKA.....	5
2.1 Studi Literatur.....	5
2.2 Dasar Teori.....	12
2.2.1 Anies Baswedan.....	12
2.2.2 Twitter.....	12
2.2.3 Sentimen Analisis.....	13
2.2.4 <i>Text Mining</i> .....	14
2.2.5 <i>Pre-Processing</i> .....	15
2.2.6 <i>TextBlob</i> .....	17
2.2.7 Klasifikasi.....	17
2.2.8 <i>Support Vector Machine</i> .....	18
2.2.9 <i>Support Vector Machine (SVM) For Multiclass</i> .....	23

2.2.10	<i>K-Fold Cross Validation</i> .....	24
2.2.11	<i>Confussion Matrix</i> .....	25
2.2.12	<i>Word Cloud</i> .....	27
<b>BAB III METODE PENELITIAN</b> .....		28
3.1	Objek Penelitian.....	28
3.2	Alur Penelitian.....	28
3.2.1	Pengumpulan Data.....	29
3.2.2	<i>Preprocessing Data</i> .....	29
3.2.3	<i>Translate Dataset</i> .....	32
3.2.4	<i>Labeling Dataset</i> .....	33
3.2.5	Pembobotan Kata.....	33
3.2.6	Pembagian Data Uji dan Data Latih.....	33
3.2.7	Metode Klasifikasi.....	34
3.2.8	Evaluasi Model.....	34
3.3	Alat dan Bahan.....	34
<b>BAB IV HASIL DAN PEMBAHASAN</b> .....		36
4.1	Pengumpulan Data.....	36
4.2	<i>Support Vector Machine</i> .....	38
4.2.1	Skenario Pengujian.....	39
4.3	Diskusi Hasil.....	45
<b>BAB V PENUTUP</b> .....		47
5.1	Kesimpulan.....	47
5.2	Saran.....	47
<b>REFERENSI</b> .....		48
<b>LAMPIRAN</b> .....		54

## DAFTAR TABEL

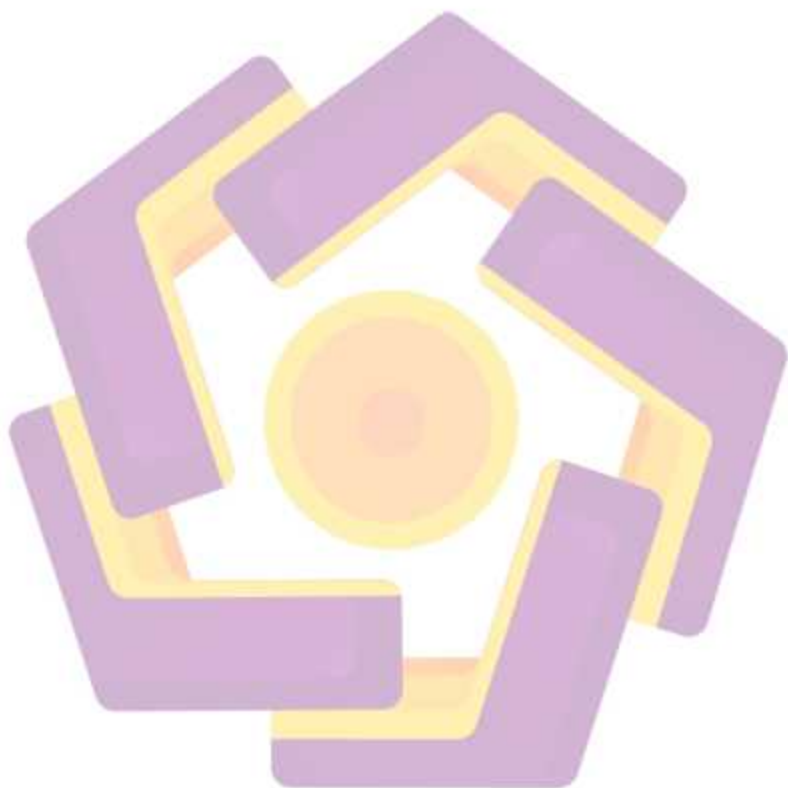
Tabel 2.1 Keaslian Penelitian .....	8
Tabel 2.2 <i>Confusion matrix</i> 2x2.....	25
Tabel 2.3 <i>Confusion matrix</i> 3x3.....	25
Tabel 3.1 <i>Case Folding</i> .....	30
Tabel 3.2 <i>Cleansing Data</i> .....	30
Tabel 3.3 <i>Tokenizing</i> .....	31
Tabel 3.4 <i>Stopword Removal</i> .....	31
Tabel 3.5 <i>Normalization</i> .....	32
Tabel 3.6 <i>Stemming</i> .....	32
Tabel 3.7 <i>Translate Dataset</i> .....	33
Tabel 3.8 <i>Labeling Dataset</i> .....	33
Tabel 4.1 Nilai <i>Hyperparameter</i> .....	39
Tabel 4.2 Hasil Uji Kernel Linear.....	39
Tabel 4.3 <i>Confusion Matrix</i> Kernel Linear.....	40
Tabel 4.4 Hasil Uji Kernel Rbf.....	40
Tabel 4.5 <i>Confusion Matrix</i> Kernel Rbf.....	41
Tabel 4.6 Hasil Uji Kernel Sigmoid.....	41
Tabel 4.7 <i>Confusion Matrix</i> Kernel Sigmoid.....	42
Tabel 4.8 Hasil Uji Kernel <i>Polynomial</i> .....	42
Tabel 4.9 <i>Confusion Matrix</i> Kernel <i>Polynomial</i> .....	43
Tabel 4.10 <i>Hyperparameter</i> Terbaik Setiap Nilai Kernel.....	43
Tabel 4.11 Hasil Uji Setiap Kernel .....	44
Tabel 4.12 Hasil Uji Data Latih.....	45
Tabel 4.13 <i>Confusion Matrix</i> Pengujian Data Latih .....	45

## DAFTAR GAMBAR

Gambar 2.1 SVM berusaha untuk menemukan hyperplane terbaik yang memisahkan kedua kelas $-1$ dan $+1$ .....	18
Gambar 2.2 <i>Hyperplane</i> terbentuk diantara kelas $-1$ dan $+1$ .....	20
Gambar 2.3 Kernel Linear.....	21
Gambar 2.4 Kernel <i>Polynomial</i> .....	21
Gambar 2.5 Kernel RBF.....	22
Gambar 2.6 Kernel Sigmoid.....	23
Gambar 3.1 <i>Flowchart</i> Penelitian.....	28
Gambar 3.2 <i>Flowchart Pre-processing</i> .....	29
Gambar 4.1 Jumlah <i>Labeling</i> Dataset.....	36
Gambar 4.2 Visualisasi <i>Wordcloud</i> Sentimen Positif.....	37
Gambar 4.3 Visualisasi <i>Wordcloud</i> Sentimen Negatif.....	37
Gambar 4.4 Visualisasi <i>Wordcloud</i> Sentimen Netral.....	38
Gambar 4.5 <i>Classification Report</i> Kernel Linear.....	44
Gambar 4.6 Hasil Uji Kernel SVM.....	46

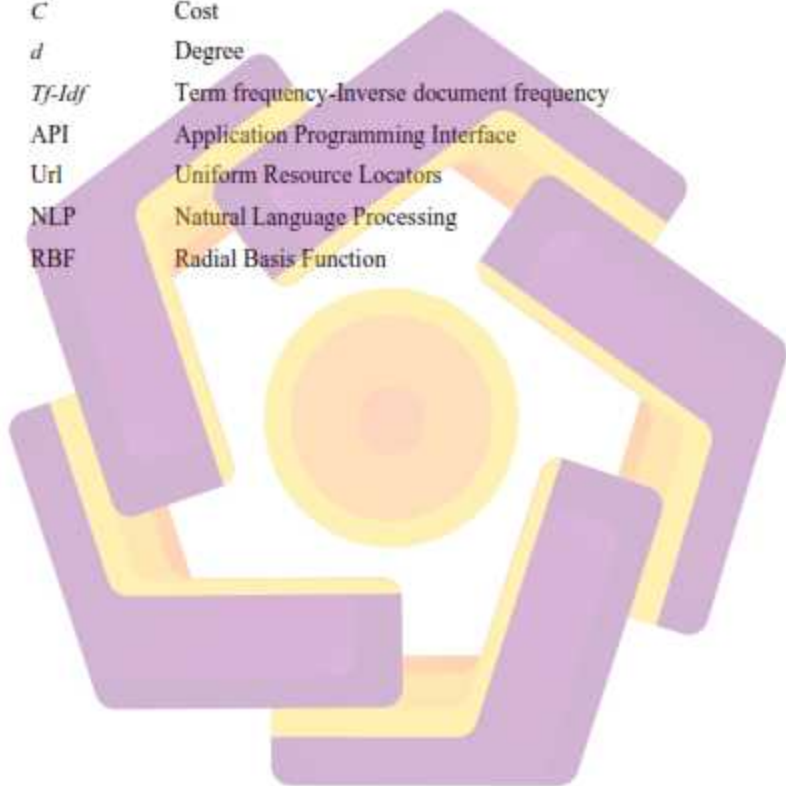
## DAFTAR LAMPIRAN

Lampiran 1. Code.....	53
-----------------------	----



## DAFTAR LAMBANG DAN SINGKATAN

SVM	Support Vector Machines
NBC	Naive Bayes
$\gamma$	Gamma
$C$	Cost
$d$	Degree
<i>Tf-Idf</i>	Term frequency-Inverse document frequency
API	Application Programming Interface
Url	Uniform Resource Locators
NLP	Natural Language Processing
RBF	Radial Basis Function





## INTISARI

Pemilihan umum 2024 semakin dekat, euforia pemilihan sudah terasa diakhir tahun 2022 ini. Partai nasional demokrat menjadi salah satu partai yang sudah mendeklarasikan calon presiden yang mereka usung. Mereka mengusung Anies Baswedan sebagai calon Presiden, hal ini tentunya menuai pro dan kontra di masyarakat. Banyak opini positif maupun negatif yang diberikan kepada Anies yang baru saja menyelesaikan jabatannya sebagai gubernur DKI Jakarta. Opini-opini tersebut banyak disampaikan melalui sosial media, salah satunya yakni Twitter. Oleh karena itu, penelitian ini mencoba memanfaatkan Twitter dengan menganalisis *tweet* dengan melakukan klasifikasi *tweet* yang berisi sentimen masyarakat terhadap Anies Baswedan. Peneliti melakukan klasifikasi menggunakan metode *Support Vector Machine (SVM)* dalam penelitian karena memiliki tingkat akurasi yang baik. Untuk mempercepat pelabelan, peneliti menggunakan metode berbasis leksikon, agar dapat dilakukan secara otomatis. Peneliti diawali dengan mengumpulkan data dari Twitter, memproses data mentah menjadi data siap pakai, melabeli data, pembobotan dengan Tf-Idf, klasifikasi menggunakan SVM dengan menguji beberapa jenis kernel dan *hyperparameter*, dan proses evaluasi menggunakan *confusion matrix* untuk menemukan performa terbaik. Hasil penelitian menunjukkan bahwa performa algoritma SVM bekerja secara optimal, dengan mendapatkan nilai akurasi sebesar 94,04%, nilai *recall* 90,21%, dan nilai *F1-score* 91,87% dengan fungsi kernel linear. Pengujian lain dengan kernel rbf mencapai nilai presisi tertinggi sebesar 93,99%.

**Kata kunci:** Sentimen Analisis, *Support Vector Machine*, *Tweet*



## ABSTRACT

The 2024 general election is getting closer, the euphoria of the election has been felt at the end of 2022. Partai Nasional Demokrat is one of the parties that has declared its presidential candidate. They carry Anies Baswedan as a presidential candidate, this of course reaps pros and cons in society. There are positive and negative opinions were given to Anies who had just finished his term as governor of DKI Jakarta. Many of these opinions are conveyed through social media, one of which is Twitter. Therefore, this study attempts to utilize Twitter by analyzing tweets by classifying tweets containing public sentiment towards Anies Baswedan. Researchers classify using the Support Vector Machine (SVM) method in research because it has a good level of accuracy. To speed up labeling, researchers use a lexicon-based method, so that it can be done automatically. Researchers collect data from Twitter, process raw data into ready-to-use data, label data, weight with Tf-Idf, classify using SVM by testing several types of kernels and hyperparameters, and the evaluation process uses a confusion matrix to find the best performance. The results showed that the performance of the SVM algorithm worked optimally, by obtaining an accuracy value of 94.04%, a recall value of 90.21%, and an F1-score value of 91.87% with a linear kernel function. Another test with the rbf kernel achieves the highest precision value of 93.99%.

**Keyword:** Sentiment Analysis, Support Vector Machine, Tweet