

**SENTIMEN ANALISIS ANTUSIASME MASYARAKAT  
TERHADAP PEMILU PRESIDEN 2024 MENGGUNAKAN  
SUPPORT VECTOR MACHINE**

**SKRIPSI**

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



disusun oleh

**TRI DAMARYANA**

**19.11.3205**

Kepada

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

**2023**

**SENTIMEN ANALISIS ANTUSIASME MASYARAKAT  
TERHADAP PEMILU PRESIDEN 2024 MENGGUNAKAN  
SUPPORT VECTOR MACHINE**

**SKRIPSI**

untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi Informatika



disusun oleh

**TRI DAMARYANA**

**19.11.3205**

Kepada

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA  
2023**

**HALAMAN PERSETUJUAN  
SKRIPSI**

**SENTIMEN ANALISIS ANTUSIASME MASYARAKAT TERHADAP PEMILU  
PRESIDEN 2024 MENGGUNAKAN SUPPORT VECTOR MACHINE**

yang disusun dan diajukan oleh

**Tri Damaryana**

19.11.3205

telah disetujui oleh Dosen Pembimbing Skripsi  
pada tanggal 7 Agustus 2023

**Dosen Pembimbing,**



**(Norhikmah, M.Kom)**

**NIK. 190302245**

**HALAMAN PENGESAHAN**

**SKRIPSI**

**SENTIMEN ANALISIS ANTUSIASME MASYARAKAT TERHADAP PEMILU  
PRESIDEN 2024 MENGGUNAKAN SUPPORT VECTOR MACHINE**

yang disusun dan diajukan oleh

**Tri Damaryana**

**19.11.3205**

Telah dipertalihkan di depan Dewan Penguji  
pada tanggal 7 Agustus 2023

**Susunan Dewan Penguji**

**Nama Penguji**

**Tanda Tangan**

**Dina Maulina, M.Kom**

**NIK. 190302250**

**Mulia Sulistiyono, M.Kom**

**NIK. 190302248**

**Norhikmah, M.Kom**

**NIK. 190302245**

Skrripsi ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Sarjana Komputer  
Tanggal 7 Agustus 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 : Tri Damaryana  
NIM : 19.11.3205

Menyatakan bahwa Skripsi dengan judul berikut:

### Sentimen Analisis Antusiasme Masyarakat Terhadap Pemilu Presiden 2024 Menggunakan Support Vector Machine

Dosen Pembimbing : Norhikmah, 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, 7 Agustus 2023

Yang Menyatakan,



Tri damaryana

## MOTTO

“Jangan menyerah menghadapi suatu masalah”

“Jawaban dari sebuah keberhasilan adalah terus belajar dan tak kenal putus asa”

“Manusia asalnya dari tanah, makan hasil tanah, berdiri di atas tanah, akan kembali ke tanah. Kenapa masih bersifat langit?” - Hamka



## HALAMAN PERSEMBAHAN

Puji dan Syukur penulis panjatkan kepada Tuhan Yang Maha Esa atas anugerah dan nikmat yang tak terkira sehingga penulis dapat menyelesaikan karya tulis ini. Pada kesempatan ini penulis ingin menyampaikan ucapan terimakasih yang sebesar-besarnya kepada :

1. Orang tua saya yang selalu memberikan dukungan, doa, dan cinta tanpa henti
2. Keluarga besar penulis yang selalu menjadi sumber inspirasi dan semangat
3. Dosen dan staf pengajar di Program Studi Teknik Informatika Universitas Amikom Yogyakarta yang telah memberikan ilmu dan pengalaman berharga selama masa studi
4. Teman-teman seangkatan dan seluruh mahasiswa Universitas Amikom yogyakarta yang selalu memberikan dukungan dan motivasi selama perkuliahan dan penyusunan skripsi ini. Penulis mengucapkan terima kasih yang sebesar-besarnya atas dukungan, doa, dan semangat yang telah diberikan. Semoga Allah SWT senantiasa memberikan keberkahan dan kesejahteraan kepada kita semua.



## KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Allah SWT yang telah memberikan rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan penyusunan skripsi dengan judul "Sentimen Analisis Antusiasme Masyarakat Terhadap Pemilu Presiden 2024 Menggunakan Support Vector Machine". Skripsi ini disusun sebagai salah satu syarat untuk menyelesaikan studi di Program Sarjana Universitas Amikom Yogyakarta.

Dengan selesainya skripsi ini, maka penulis tidak lupa mengucapkan terimakasih kepada :

1. Bapak Prof. Dr. M. Suyanto, MM selaku Rektor Universitas Amikom Yogyakarta.
2. Hanif Al Fatta, S.Kom., M.Kom, Selaku Dekan Fakultas Ilmu Komputer.
3. Norhikmah, M.Kom selaku dosen pembimbing yang telah banyak memberikan pengarahan bagi penulis dalam pembuatan skripsi.
4. Bapak-Ibu Dosen Universitas Amikom Yogyakarta, atas ilmu dan pengetahuan yang telah diberikan kepada penulis selama duduk di bangku kuliah.
5. Kedua orang tua saya, Bapak Milana dan Ibu Marsinem yang tak pernah berhenti untuk selalu memberikan selamat serta doa.
6. Seluruh teman-teman angkatan 2019, terutama untuk kelas IF 10, terimakasih atas kerjasamanya selama ini.
7. Teman-teman seperjuangan (ilham, fahmudin, filah, bagas, uji, pandu, anton, adli, dan yohan) yang telah memberikan dukungan dan terimakasih atas momen momen yang tak pernah terlupakan.

Penulis tentunya menyadari bahwa pembuatan skripsi ini masih banyak sekali kekurangan-kekurangan dan kelemahan-kelemahannya. Oleh karena itu penulis berharap kepada semua pihak agar dapat menyampaikan kritik dan saran yang membangun untuk menambah kesempurnaan skripsi ini. Namun penulis tetap berharap skripsi ini akan bermanfaat bagi semua pihak yang membacanya.

Yogyakarta, 7 Agustus 2023

Penyusun

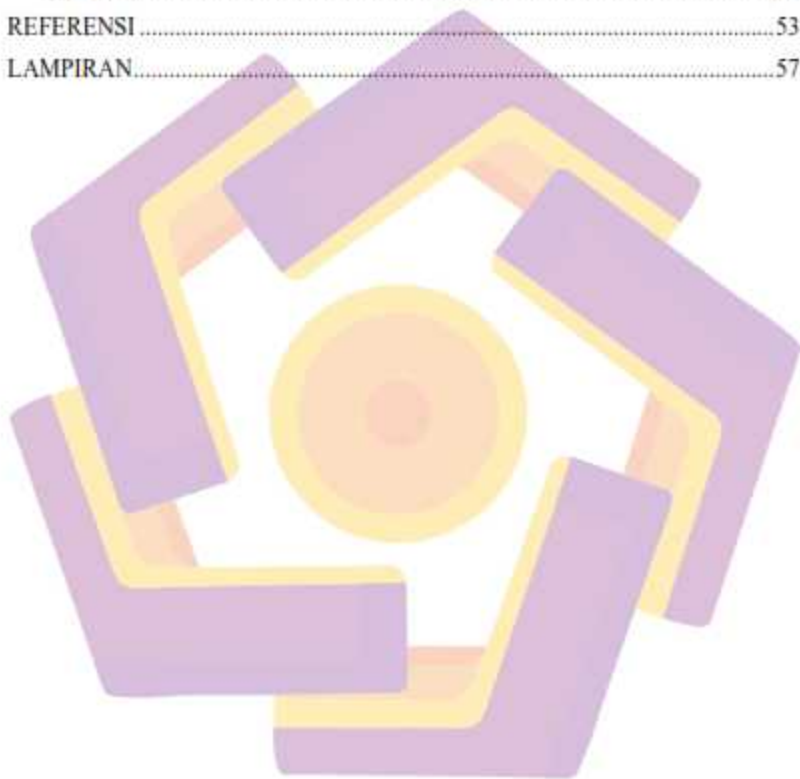


## DAFTAR ISI

HALAMAN JUDUL .....	i
HALAMAN PERSETUJUAN .....	ii
HALAMAN PENGESAHAN .....	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI .....	iv
MOTTO .....	v
HALAMAN PERSEMBAHAN .....	vi
KATA PENGANTAR .....	vii
DAFTAR ISI .....	viii
DAFTAR TABEL .....	xi
DAFTAR GAMBAR .....	xii
DAFTAR LAMPIRAN .....	xiii
DAFTAR LAMBANG DAN SINGKATAN .....	xiv
DAFTAR ISTILAH .....	xv
INTISARI .....	xvi
ABSTRACT .....	xvii
BAB I PENDAHULUAN .....	1
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	2
1.3 Batasan Masalah .....	2
1.4 Tujuan Penelitian .....	2
1.5 Manfaat Penelitian .....	3
1.6 Sistematika Penulisan .....	3
BAB II TINJAUAN PUSTAKA .....	5
2.1 Studi Literatur .....	5
2.2 Dasar Teori .....	10
2.2.1 Sentimen Analisa .....	10
2.2.2 Pemilihan Umum Presiden .....	10
2.2.3 Crawling Data .....	11
2.2.3 Youtube .....	11
2.2.4 Klasifikasi .....	12

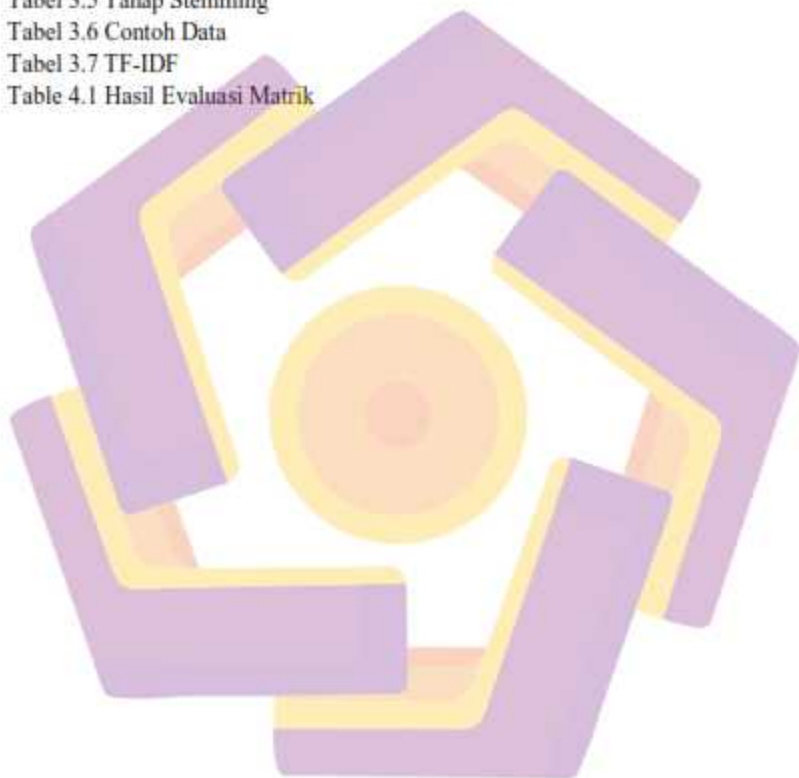
2.2.5	<i>Pre-processing</i> .....	12
2.2.6	Text Mining .....	13
2.2.7	Pembobotan TF-IDF .....	14
2.2.8	SMOTE (Synthetic Minority Over-sampling Technique) .....	14
2.2.9	Support Vector Machine .....	15
2.2.10	Confusion Matrix .....	17
2.2.11	<i>Supervised Learning</i> .....	19
2.2.12	Jupyter Notebook .....	19
2.2.13	Python .....	20
<b>BAB III METODOLOGI PENELITIAN</b> .....		22
3.1	Metodologi Penelitian .....	22
3.2	Alur Penelitian .....	25
3.2.1	Prosedur Penelitian .....	25
3.2.2	Praproses Data .....	26
3.2.3	Metode Analisis Data .....	29
3.2.4	Arsitektur Sistem .....	30
3.3	TF-IDF (Term Frequency - Inverse Document Frequency) .....	30
3.4	SMOTE (Synthetic Minority Oversampling Technique) .....	32
3.5	Klasifikasi <i>Support Vector Machine (SVM)</i> .....	33
3.6	Rancangan Pengujian Model .....	34
3.7	Alat dan Bahan Penelitian .....	35
3.7.2	Spesifikasi Perangkat Keras ( <i>Hardware</i> ) .....	35
3.7.3	Spesifikasi Perangkat Lunak ( <i>Software</i> ) .....	35
3.7.3	Bahan penelitian .....	36
<b>BAB IV HASIL DAN PEMBAHASAN</b> .....		38
4.1	Proses Scraping Data .....	38
4.2	<i>Text Preprocessing</i> .....	38
4.2.1	Case Folding .....	39
4.2.2	Cleaning .....	39
4.2.3	Tokenizing .....	40
4.2.4	Stopword Removal .....	41
4.2.5	Stemming .....	43
4.3	Labeling Text .....	44

4.4 Membagi Data Latih dan Data Uji .....	46
4.5 TF-IDF ( Term Frequency-Inverse Document Frequency ).....	46
4.6 SMOTE ( Synthetic Minority Over-sampling Technique) .....	47
4.7 Hasil Klasifikasi Sentimen .....	47
BAB V PENUTUP .....	52
5.1 Kesimpulan .....	52
5.2 Saran .....	52
REFERENSI .....	53
LAMPIRAN.....	57



## DAFTAR TABEL

- Tabel 2.1 keaslian penelitian
- Tabel 2.2 Confusion matrix
- Tabel 3.1 Tahap Case Folding
- Tabel 3.2 Tahap Clearing
- Tabel 3.3 Tahap Tokenizing
- Tabel 3.4 Tahap Stopword Removal
- Tabel 3.5 Tahap Stemming
- Tabel 3.6 Contoh Data
- Tabel 3.7 TF-IDF
- Table 4.1 Hasil Evaluasi Matrik

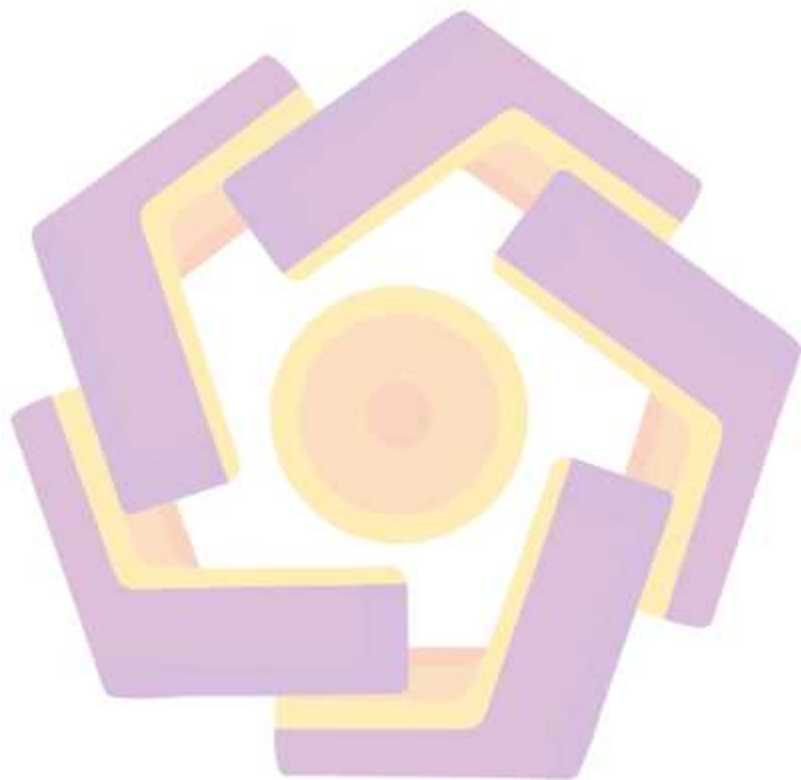


## DAFTAR GAMBAR

- Gambar 2.1. *preprocessing* data  
Gambar 2.2. *hyperplane* positif dan negatif  
Gambar 3.1 Prosedur Penelitian  
Gambar 3.2 Preprocessing data  
Gambar 3.3 Arsitektur Sistem  
Gambar 3.4 SMOTE  
Gambar 3.5 Menggunakan SMOTE  
Gambar 3.6 Tanpa SMOTE  
Gambar 3.7 Kernel Linear  
Gambar 4.1 Dataset  
Gambar 4.2 Code Program Case Folding  
Gambar 4.3 Hasil Case Folding  
Gambar 4.4 Code Program Cleaning  
Gambar 4.5 Hasil Cleaning  
Gambar 4.6 Code Program Tokenizing  
Gambar 4.7 Hasil Tokenizing  
Gambar 4.8 Code Program Stopword Removal  
Gambar 4.9 Hasil Stopword Removal  
Gambar 4.10 Code Program Stemming  
Gambar 4.11 Hasil Stemming  
Gambar 4.12 Code Program TextBlob  
Gambar 4.13 Code Program Data Latih dan Data Uji  
Gambar 4.14 Code Program TF-IDF  
Gambar 4.15 Jumlah Sebelum SMOTE  
Gambar 4.16 Jumlah Data Setelah SMOTE  
Gambar 4.17 Hasil percobaan pertama  
Gambar 4.18 Hasil percobaan kedua  
Gambar 4.19 Hasil percobaan ketiga  
Gambar 4.20 Hasil percobaan keempat  
Gambar 4.21 Hasil percobaan kelima  
Gambar 4.22 Hasil percobaan keenam

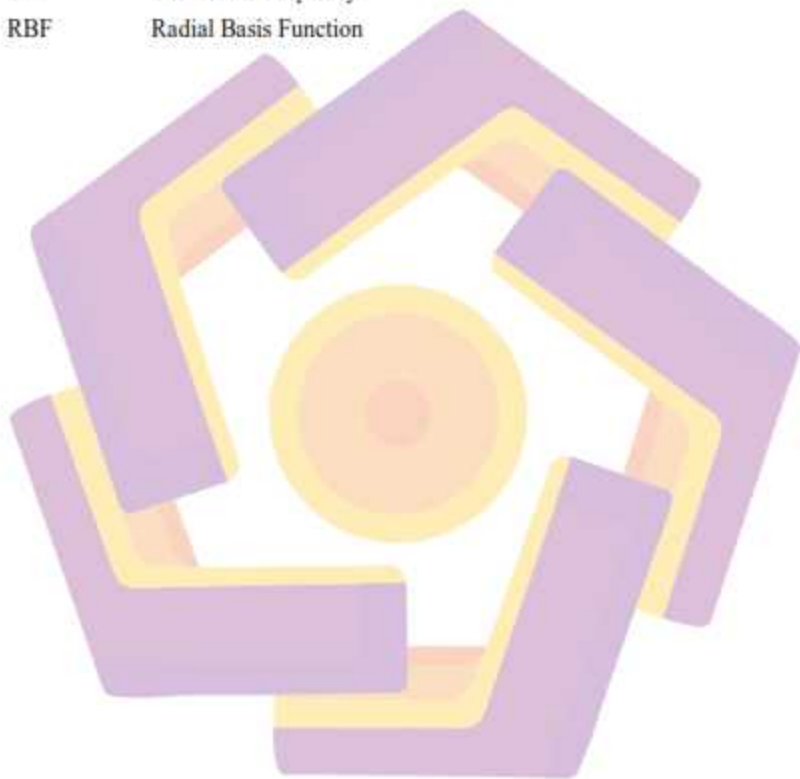
## DAFTAR LAMPIRAN

Code Crawling Data Youtube



## DAFTAR LAMBANG DAN SINGKATAN

SVM	Support Vector Machines
SMOTE	Synthetic Minority Oversampling Technique.
TF-IDF	Term Frequency - Inverse Document Frequency
N	Jumlah Total Data
DF	Document Frequency
RBF	Radial Basis Function



## DAFTAR ISTILAH

### **Scraping**

Teknik pengumpulan informasi dari situs web secara otomatis dengan menggunakan perangkat lunak atau skrip khusus.

### **Hyperplane**

Garis atau permukaan linear yang digunakan untuk memisahkan dua kelas dalam masalah klasifikasi





## INTISARI

Kemajuan ilmu pengetahuan yang dicirikan oleh munculnya teknologi modern menyebabkan bertambahnya unggahan pada media digital sehingga dengan bertambahnya unggahan pada media digital menyebabkan banyaknya data dari opini masyarakat pada media sosial seperti youtube. Penelitian ini bertujuan untuk melakukan Sentimen Analisis terhadap antusiasme masyarakat terhadap Pemilu Presiden 2024 menggunakan metode Support Vector Machine (SVM) dengan dan tanpa menggunakan teknik SMOTE (Synthetic Minority Over-sampling Technique). Sentimen analisis dilakukan untuk mengidentifikasi apakah masyarakat menyampaikan pandangan positif atau negatif terkait Pemilu Presiden 2024 dengan data yang diperoleh dari platform youtube. Pengujian model dilakukan dengan pembagian data uji dan data latih yang berbeda. Hasil percobaan menunjukkan bahwa penggunaan metode SVM menghasilkan tingkat akurasi rata-rata lebih dari 85%. Namun, pada pengujian menggunakan SMOTE, ditemukan bahwa nilai recall untuk kelas minoritas (Negative) cenderung lebih tinggi, menunjukkan bahwa model lebih baik dalam mengenali sentimen negatif dari masyarakat. Namun, penggunaan SMOTE juga menyebabkan nilai precision untuk kelas minoritas menjadi lebih rendah.

Kata kunci : Sentimen analisis, pemilu presiden 2024, Support Vector Machine (SVM), SMOTE (Synthetic Minority Over-sampling Technique), confusion matrix

## ABSTRACT

*The advancement of science which is characterized by the emergence of modern technology has led to an increase in uploads on digital media so that with the increase in uploads on digital media it has caused a large amount of data from public opinion on social media such as YouTube. This study aims to conduct a Sentiment Analysis of public enthusiasm for the 2024 Presidential Election using the Support Vector Machine (SVM) method with and without using the SMOTE (Synthetic Minority Over-sampling Technique) technique. Sentiment analysis was carried out to identify whether the public expressed positive or negative views regarding the 2024 Presidential Election using data obtained from the YouTube platform. Model testing is done by dividing the different test data and training data. The experimental results show that the use of the SVM method produces an average accuracy rate of more than 85%. However, in testing using SMOTE, it was found that the recall value for the minority class (Negative) tends to be higher, indicating that the model is better at recognizing negative sentiments from the public. However, the use of SMOTE also causes the precision value for the minority class to be lower.*

**Keywords:** *Sentiment analysis, 2024 presidential election, Support Vector Machine (SVM), SMOTE (Synthetic Minority Over-sampling Technique), confusion matrix*