

**ANALISIS SENTIMEN TWITTER MENGGUNAKAN  
METODE NAÏVE BAYES DAN SVM TERHADAP GUBERNUR  
JAWA TENGAH**

**SKRIPSI**

Diajukan untuk memenuhi salah satu syarat m encapai derajat Sarjana  
Program Studi Informatika



disusun oleh

**MUKHAMAD YUSUF AFANDI**

**19.11.2803**

Kepada

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

**ANALISIS SENTIMEN TWITTER MENGGUNAKAN  
METODE NAÏVE BAYES DAN SVM TERHADAP GUBERNUR  
JAWA TENGAH**

**SKRIPSI**

untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi Informatika



disusun oleh

**MUKHAMAD YUSUF AFANDI**

**19.11.2803**

Kepada

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

**2023**

**HALAMAN PERSETUJUAN**

**SKRIPSI**

**ANALISIS SENTIMEN TWITTER MENGGUNAKAN  
METODE NAÏVE BAYES DAN SVM TERHADAP GUBERNUR  
JAWA TENGAH**

yang disusun dan diajukan oleh

**MUKHAMAD YUSUF AFANDI**

**19.11.2803**

telah disetujui oleh Dosen Pembimbing Skripsi  
pada tanggal 24 Juli 2023

Dosen Pembimbing,



**Arif Dwi Laksito, M. Kom**

**NIK. 190302150**

**HALAMAN PENGESAHAN**

**SKRIPSI**

**ANALISIS SENTIMEN TWITTER MENGGUNAKAN  
METODE NAÏVE BAYES DAN SVM TERHADAP GUBERNUR  
JAWA TENGAH**

yang disusun dan diajukan oleh

**MUKHAMAD YUSUF AFANDI**

**19.11.2803**

Telah dipertahankan di depan Dewan Penguji  
pada tanggal 24 Juli 2023

**Susunan Dewan Penguji**

**Nama Penguji**

**Tanda Tangan**

**Ike Verawati, M.Kom**  
**NIK. 190302237**

**Eli Pujastuti, M.Kom**  
**NIK. 190302227**

**Arif Dwi Laksito, M.Kom**  
**NIK. 190302150**



Skripsi ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Sarjana Komputer  
Tanggal 24 Juli 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 : MUKHAMAD YUSUF AFANDI**  
**NIM : 19.11.2803**

Menyatakan bahwa Skripsi dengan judul berikut:

### **ANALISIS SENTIMEN TWITTER MENGGUNAKAN METODE NAÏVE BAYES DAN SVM TERHADAP GUBERNUR JAWA TENGAH**

Dosen Pembimbing : Arif Dwi Laksito, 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, 24 Juli 2023

Yang Menyatakan,



Mukhamad Yusuf Afandi

## HALAMAN PERSEMBAHAN

Dengan segala puji syukur kepada Allah SWT dan atas dukungan dan do'a dari orang tua tecinta, akhirnya Skripsi ini dapat di selesaikan dengan baik dan tepat waktu. Oleh karena itu, dengan rasa bangga dan bahagia saya ucapkan rasasyukur dan terima kasih kepada.

1. Allah SWT, karena hanya atas izin dan karunianyaalah maka skripsi ini dapat di buat dan selesai pada waktunya.
2. Bapak H. Sofyan dan ibu Soimatun Hj. Yang telah memberikan dukungan moril dan material tidak lupa dengan doa yang selalu menyertai ku tiada henti untuk kesuksesan dan keberhasilan sampai sekarang ini, sebab tidak ada kata indah selain doa dan tiada doa yang paling khusuk kecuali doa dari orang tua.
3. Adik - adik saya tercinta yang selalu memberi do'a dan semangat untuk menyelesaikan skripsi ini.
4. Para sahabat dan semua pihak yang telah membantu dalam pelaksanaan penelitian skripsi ini hingga ujian.

## KATA PENGANTAR

Alhamdulillah, puji dan syukur penulis panjatkan kehadiran Allah SWT karena atas limpahan dan rahmat, ridho dan karunia-Nya penulis masih diberi kesempatan dan kemudahan untuk menyelesaikan skripsi ini dengan judul “Analisis Sentimen Twitter Menggunakan Metode Naïve Bayes Dan Svm Terhadap Gubernur Jawa Tengah” sesuai dengan yang diharapkan. Skripsi ini penulis persembahkan teruntuk:

1. Bapak Prof. Dr. M. Suyanto, MM selaku rektor Universitas Amikom Yogyakarta.
2. Bapak Arif Dwi Laksito, M. Kom selaku dosen pembimbing saya yang telah sudah memberikan bayak arahan dan bimbingan dalam pelaksanaan skripsi ini.
3. Bapak, Ibu dosen, seluruh staf dan pegawai di prodi Informatika yang telah membimbing dan menjadi bagian pembelajaran diri selama studi.
4. Ayah dan ibu tercinta dan tersayang yang telah memberikan dorongan moral maupun material selama studidan menyelesaikan skripsi ini.
5. Serta semua pihak yang telah membantu dan bekerja sama dalam pelaksanaan skripsi ini.

Penulis menyadari masih mempunyai banyak kekurangan dalam penyusun laporan skripsi ini. Maka dari itu kritik dan saran adalah hal yang sangat kami harapkan demi dapat kemajuan bersama dan meningkatkan ilmu pengetahuan di Indonesia.

Yogyakarta, 07 Juli 2023

Penulis

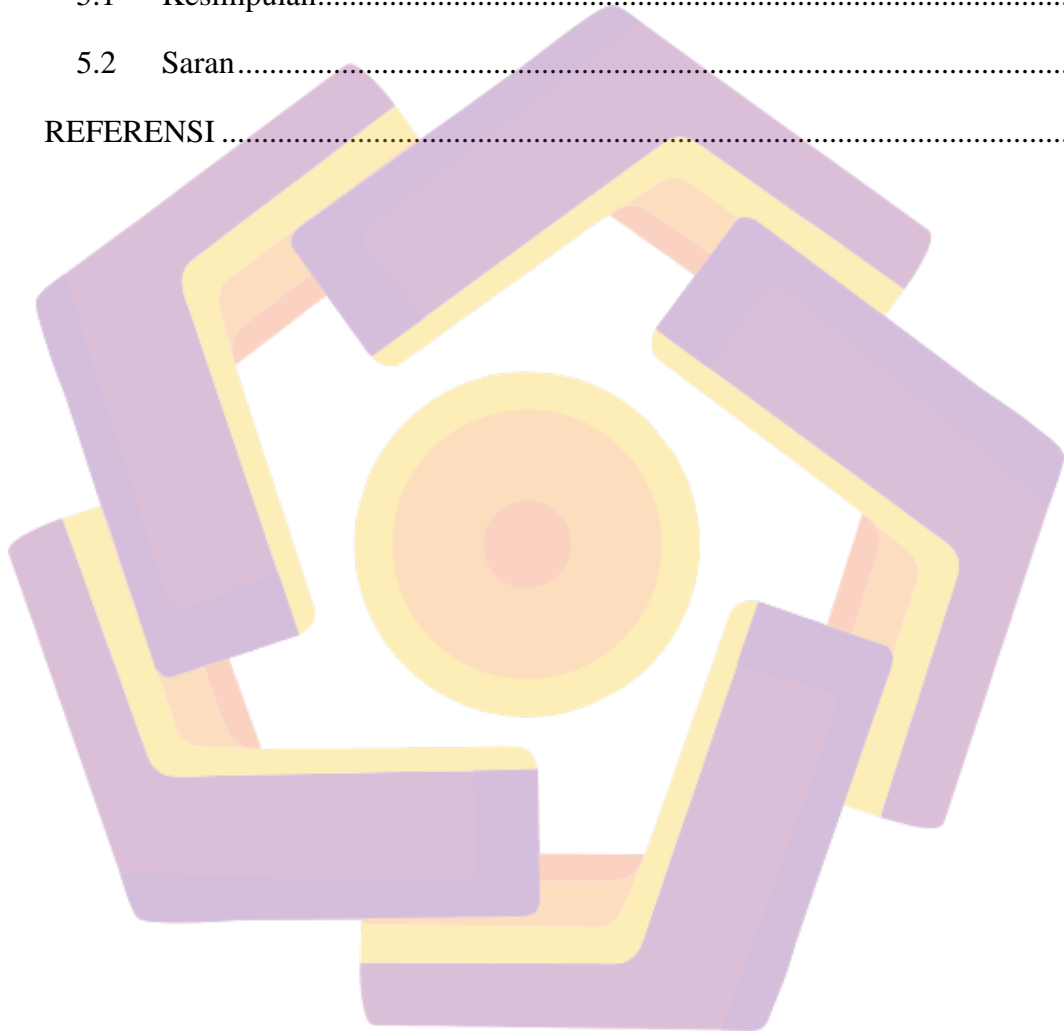
## DAFTAR ISI

HALAMAN JUDUL .....	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN .....	iii
HALAMAN PERSEMBAHAN .....	v
KATA PENGANTAR .....	vi
DAFTAR ISI.....	vii
DAFTAR TABEL.....	x
DAFTAR GAMBAR .....	xi
DAFTAR LAMBANG DAN SINGKATAN .....	xii
DAFTAR ISTILAH.....	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.....	17
2.2.1 Analisis Sentimen .....	17



2.2.2	Machine Learning .....	17
2.2.3	Supervised Learning .....	18
2.2.4	Unsupervised Learning .....	18
2.2.5	Reinforcement Learning .....	19
2.2.6	Twitter.....	19
2.2.7	Text Mining .....	20
2.2.8	Naïve Bayes .....	21
2.2.9	Support Vector Machine .....	22
2.2.10	Confusion Matrix .....	23
2.2.11	Python .....	25
2.2.12	Jupyter Notebook .....	26
<b>BAB III METODE PENELITIAN .....</b>		<b>27</b>
3.1	Objek Penelitian .....	27
3.2	Alur Penelitian.....	27
3.2.1	Data Collection .....	28
3.2.2	Preprocessing .....	28
3.2.3	Ekstraksi Fitur .....	28
3.2.4	Model Data.....	29
3.2.5	Evaluasi.....	30
3.3	Alat dan Bahan .....	30
<b>BAB IV HASIL DAN PEMBAHASAN .....</b>		<b>32</b>
4.1	Hasil dan Pembahasan .....	32
4.1.1	Crawling Data .....	32
4.1.2	Preprocessing .....	33
4.1.3	TF-IDF .....	35

4.1.4 Oversampling.....	38
4.1.5 Model Data.....	40
4.1.6 Evaluasi.....	41
<b>BAB V PENUTUP .....</b>	<b>45</b>
5.1 Kesimpulan.....	45
5.2 Saran.....	46
<b>REFERENSI .....</b>	<b>47</b>



## DAFTAR TABEL

<b>Tabel 2. 1</b> Keaslian Penelitian	12
<b>Tabel 2. 2</b> Confusion Matrix	24
<b>Tabel 3. 1</b> Spesifikasi perangkat keras	31
<b>Tabel 4. 1</b> Preprocessing	34
<b>Tabel 4. 2</b> Labeling data	35
<b>Tabel 4. 3</b> Tweet dan Label	36
<b>Tabel 4. 4</b> Menghitung Term Frequency	36
<b>Tabel 4. 5</b> Menghitung Inverse Document Frequency	37
<b>Tabel 4. 6</b> Menghitung TF-IDF	38
<b>Tabel 4. 7</b> Label dan Jumlah data	39
<b>Tabel 4. 8</b> Label dan Jumlah data	40
<b>Tabel 4. 9</b> hasil <i>accuracy</i> , <i>recall</i> dan <i>F1-Score</i> pada metode Naïve Bayes	43
<b>Tabel 4. 10</b> hasil <i>accuracy</i> , <i>recall</i> dan <i>F1-Score</i> pada metode SVM	44

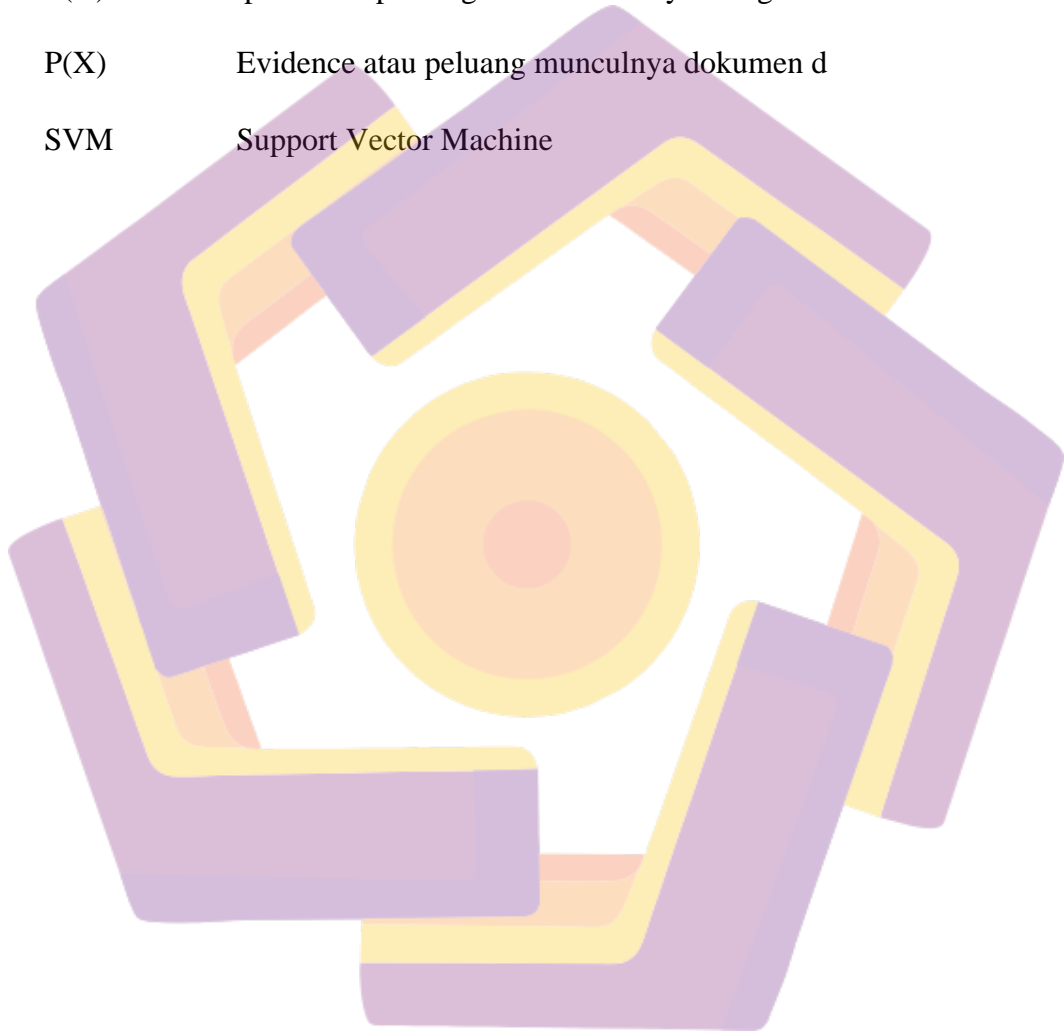
## DAFTAR GAMBAR

<b>Gambar 3. 1</b> Alur Penelitian	27
<b>Gambar 4. 1</b> Contoh Tweet	32
<b>Gambar 4. 2</b> Hasil crawling data Tweet	33
<b>Gambar 4. 3</b> Barchart Sebelum Oversampling	39
<b>Gambar 4. 4</b> Barchart Sesudah Oversampling	40
<b>Gambar 4. 5</b> hasil Confusion Matrix Model Naïve Bayes	42
<b>Gambar 4. 6</b> hasil Confusion Matrix Model SVM	43



## DAFTAR LAMBANG DAN SINGKATAN

$P(H X)$	posterior yaitu peluang kelas H dan diberikan dokumen D
$P(X H)$	nilai Likelihood
$P(H)$	prior atau peluang awal munculnya kategori H
$P(X)$	Evidence atau peluang munculnya dokumen d
SVM	Support Vector Machine



## DAFTAR ISTILAH

Vektor	Besaran yang mempunyai arah
Platform	Perangkat untuk menjalankan Hardware Software
Likelihood	Peluang atau kemungkinan
Dataset	Himpunan data
Tweet	pesan yang dikirim oleh pengguna berupa teks, tautan, gambar dan video



## INTISARI

Analisis sentimen adalah salah satu bidang dari *Natural Language Processing*(NLP). Suatu metode yang digunakan untuk mengidentifikasi dan mengevaluasi isi dari dataset dengan opini atau pandangan. Sehingga banyak digunakan untuk kepentingan sentimen dalam suatu objek. Pada penelitian ini yaitu untuk mengetahui sentimen publik media sosial twitter dengan menggunakan kata kunci di pencarian Ganjar Pranowo dikarenakan banyak dukungan dari masyarakat sebagai calon presiden 2024. Selain mendapatkan dukungan Ganjar Pranowo juga mendapatkan berbagai kritik dan saran bahkan ujaran kebencian.

Metode yang digunakan pada analisis sentimen ini adalah perbandingan algoritma Naïve Bayes dan Support Vector Machine (SVM). Tahapan penelitian ini mengambil data dengan menggunakan metode crawling tweet dengan total 1229 data. Data tweet yang telah di crawling kemudian dilakukan preprocessing untuk di bersihkan lalu dihapus data yang tidak diperlukan dan melakukan metode ekstraksi fitur. Selanjutnya data akan diproses melalui Klasifikasi sentimen menggunakan perbandingan pembelajaran mesin (machine learning) seperti Naïve Bayes dan Support Vector Machine .

Hasil dari penelitian menunjukkan bahwa Support Vector Machine mendapatkan kinerja yang lebih baik dengan nilai accuracy 94% sedangkan untuk Naïve Bayes mendapatkan hasil accuracy 87%.

**Kata kunci:** Sentimen analisis, Klasifikasi, Naïve bayes dan Support Vector Machine, Tweet.

## ABSTRACT

*Sentiment analysis is a field of Natural Language Processing(NLP). A method used to identify and evaluate the contents of a dataset with opinions or views. So that it is widely used for the sake of sentiment in an object. In this study, it is to find out public sentiment on Twitter social media by using keywords in the search for Ganjar Pranowo because there is a lot of support from the public as a 2024 presidential candidate. Apart from getting support for Ganjar Pranowo, he also gets various criticisms and suggestions and even hate speech.*

*The method used in this sentiment analysis is a comparison of the Naïve Bayes algorithm and the Support Vector Machine (SVM). The stage of this research was to collect data using the tweet crawling method with a total of 1229 data. The tweet data that has been crawled is then pre-processed to be cleaned and then deleted data that is not needed and used the extraction future. Furthermore, the data will be processed through sentiment classification using machine learning comparisons such as Naïve Bayes and Support Vector Machine.*

*The results of the research show that the Support Vector Machine gets better performance with an accuracy value of 94% while for Naïve Bayes it gets an accuracy of 87%.*

**Keyword** Sentimen analisis, Klasifikasi, Naïve bayes dan Support Vector Machine, Tweet.