

**ANALISIS SENTIMEN MASA TRANSISI PEMERINTAHAN
PRESIDEN JOKO WIDODO MENGGUNAKAN ALGORITMA
SVM DAN NAÏVE BAYES**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi S1 Informatika



disusun oleh

ALFA ZIDAN PRASTYO

21.11.4315

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

**ANALISIS SENTIMEN MASA TRANSISI PEMERINTAHAN
PRESIDEN JOKO WIDODO MENGGUNAKAN ALGORITMA
SVM DAN NAÏVE BAYES**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi S1 Informatika



disusun oleh

ALFA ZIDAN PRASTYO

21.11.4315

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

HALAMAN PERSETUJUAN

SKRIPSI

**Analisis Sentimen Masa Transisi Pemerintahan Presiden Joko
Widodo Menggunakan Algoritma SVM dan Naïve Bayes**

yang disusun dan diajukan oleh

ALFA ZIDAN PRASTYO

21.11.4315

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 13 Februari 2025

Dosen Pembimbing,



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

HALAMAN PENGESAHAN

SKRIPSI

**Analisis Sentimen Masa Transisi Pemerintahan Presiden Joko
Widodo Menggunakan Algoritma SVM dan Naïve Bayes**

yang disusun dan diajukan oleh

ALFA ZIDAN PRASTYO

21.11.4315

Telah dipertahankan di depan Dewan Penguji
pada tanggal 13 Februari 2025

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Dr. Andi Sunyoto, S.Kom., M.Kom.
NIK. 190302052

Dr. Kumara Ari Yuana, ST, MT
NIK. 190302575

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



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 13 Februari 2025

DEKAN FAKULTAS ILMU KOMPUTER



Hanif Al Fatta, S.Kom., M.Kom., Ph.D.
NIK. 190302096

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : ALFA ZIDAN PRASTYO
NIM : 21.11.4315

Menyatakan bahwa Skripsi dengan judul berikut:

Analisis Sentimen Masa Transisi Pemerintahan Presiden Joko Widodo Menggunakan Algoritma SVM dan Naïve 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, 13 Februari 2025

Yang Menyatakan,



ALFA ZIDAN PRASTYO

KATA PENGANTAR

Puji syukur saya panjatkan kepada Allah SWT, yang telah memberikan rahmat, hidayah, serta karunia-Nya, sehingga saya dapat menyelesaikan skripsi ini dengan baik dan tepat waktu. Skripsi yang berjudul "Analisis Sentimen Masa Transisi Pemerintahan Presiden Joko Widodo Menggunakan Algoritma SVM dan Naïve Bayes" ini diajukan sebagai salah satu syarat untuk menyelesaikan studi program strata satu (S1) pada Program Studi Informatika Universitas AMIKOM Yogyakarta. Penulis mendapatkan dukungan dari banyak pihak dalam proses pengerjaan skripsi ini. Oleh karena itu, penulis mengucapkan terima kasih kepada:

1. Bapak Majid Rahardi, S.Kom., M.Eng., selaku Dosen Pembimbing Skripsi saya, yang telah memberikan bimbingan, arahan, dan saran yang sangat berharga sepanjang proses penelitian dan penulisan skripsi ini.
2. Dosen Penguji, penulis menyampaikan rasa hormat dan ucapan terima kasih yang sebesar-besarnya atas kesediannya untuk menguji, memberikan masukan, serta arahan yang sangat berharga.
3. Untuk kedua orang tua saya yang selalu memberikan doa, kasih sayang, serta dukungan tanpa henti, yang tidak hanya memotivasi saya untuk terus berusaha, tetapi juga memberikan semangat dalam setiap langkah yang saya ambil.
4. Para Dosen dan Staf Universitas AMIKOM Yogyakarta, yang telah memberikan ilmu pengetahuan yang sangat bermanfaat bagi perkembangan pribadi saya selama masa kuliah.
5. Kepada teman-teman, yang telah memberikan dukungan moral, semangat, dan bantuan dalam menyelesaikan tugas-tugas, khususnya di waktu-waktu yang penuh tantangan, serta selalu berbagi cerita dan kebersamaan selama menjalani masa perkuliahan ini.

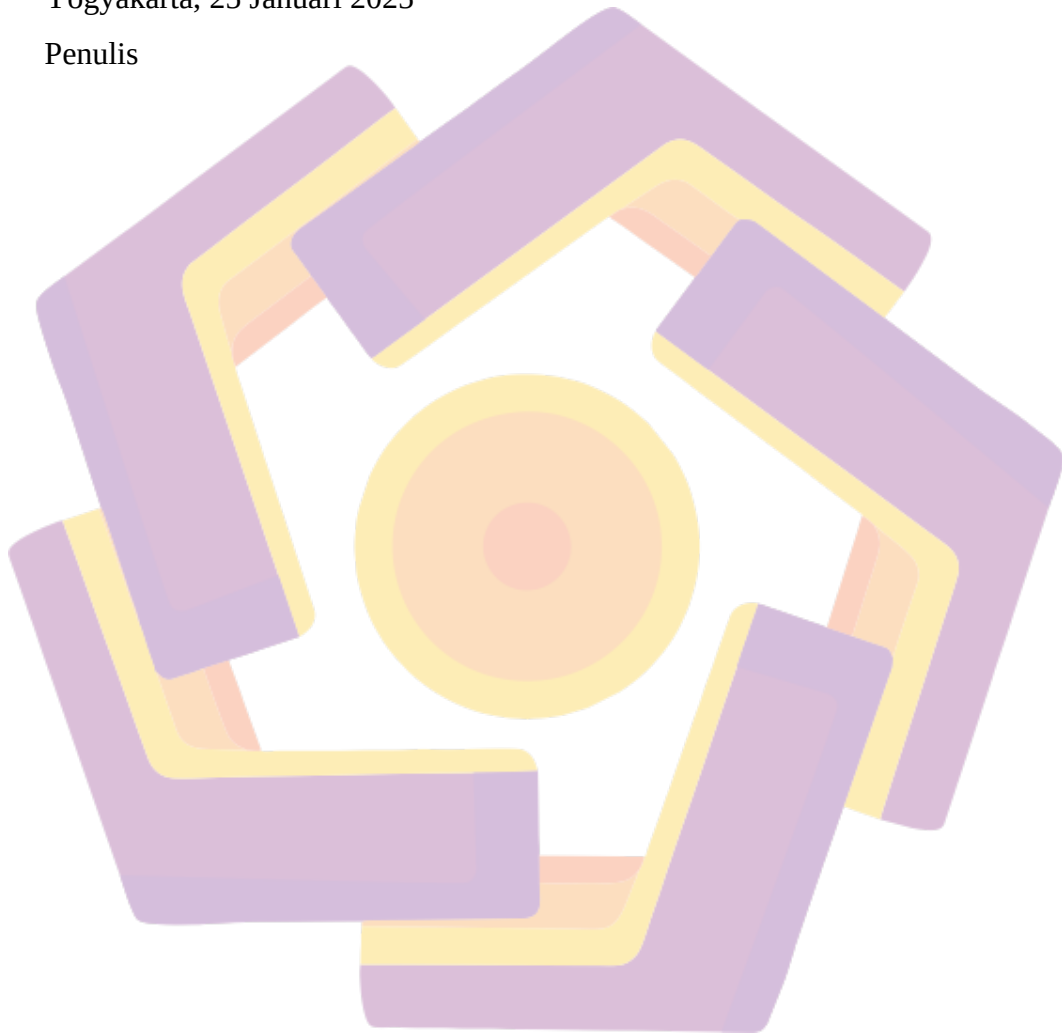
Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Oleh karena itu, kritik dan saran yang membangun sangat penulis harapkan untuk perbaikan di

masa yang akan datang. Semoga skripsi ini dapat memberikan manfaat dan kontribusi bagi perkembangan ilmu pengetahuan.

Akhir kata, penulis mengucapkan terima kasih kepada semua pihak yang telah membantu dalam penyelesaian skripsi ini.

Yogyakarta, 25 Januari 2025

Penulis



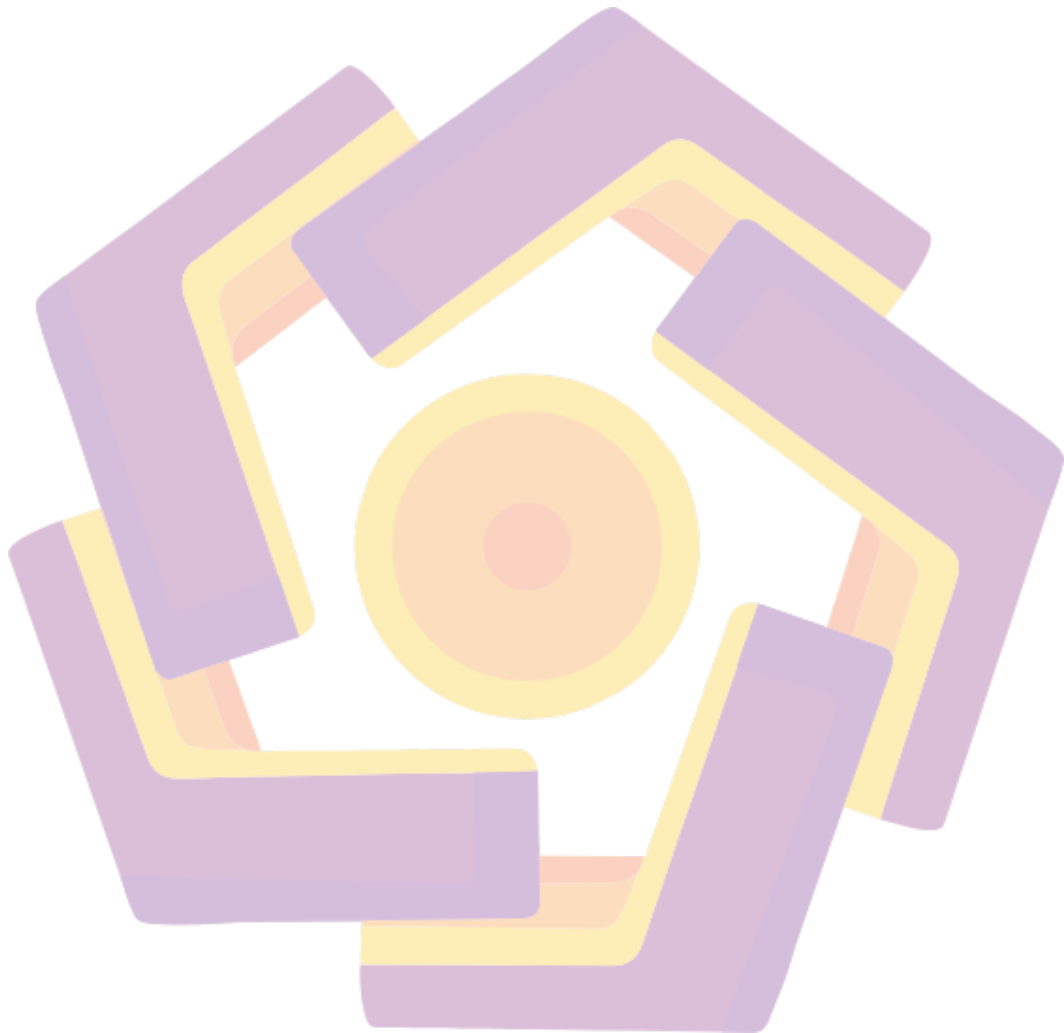
DAFTAR ISI

HALAMAN JUDUL	I
HALAMAN PERSETUJUAN.....	II
HALAMAN PENGESAHAN	III
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	IV
KATA PENGANTAR	V
DAFTAR ISI.....	VII
DAFTAR TABEL.....	XI
DAFTAR GAMBAR.....	XII
DAFTAR LAMPIRAN.....	XIV
DAFTAR LAMBANG DAN SINGKATAN	XV
INTISARI	XVI
ABSTRACT.....	XVII
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang Masalah.....	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah.....	3
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	4
1.6 Sistematika Penulisan.....	4
BAB II TINJAUAN PUSTAKA	6
2.1 Studi Literatur	6
2.2 Dasar Teori.....	13
2.2.1. Machine Learning	13

2.2.2.	Natural Language Processing	13
2.2.3.	Text Mining	13
2.2.4.	Analisis Sentimen	14
2.2.5.	YouTube	14
2.2.6.	Masa Transisi	14
2.2.7.	Preprocessing	15
2.2.8.	TextBlob	15
2.2.9.	Support Vector Machine	16
2.2.10.	Naïve Bayes.....	17
2.2.11.	Confusion Matrix	18
2.2.12.	Term Frequency - Inverse Document Frequency (TF-IDF).....	19
2.2.13.	Latent Dirichlet Allocation (LDA).....	20
2.2.14.	Flask	20
BAB III METODE PENELITIAN		21
3.1	Objek Penelitian	21
3.2	Alur Penelitian.....	21
3.2.1.	Crawling Data	22
3.2.2.	Preprocessing	23
3.2.3.	Labeling	27
3.2.4.	Pemodelan.....	28
3.2.5.	Visualisasi	32
3.2.6.	Analisa Topic Modeling (LDA).....	32
3.2.7.	Implementasi Hasil Model	33
3.3	Alat dan Bahan	34
3.3.1.	Data Penelitian	34

3.3.2. Alat Penelitian.....	35
BAB IV HASIL DAN PEMBAHASAN	36
4.1. Pengumpulan Data	36
4.2. Preprocessing	38
4.2.1. Cleaning	39
4.2.2. Case Folding	41
4.2.3. Normalisasi	42
4.2.4. Tokenize.....	43
4.2.5. Stopword Removal.....	44
4.2.6. Stemming	45
4.2.7. Data Quality Enhancement	46
4.3. Labeling.....	46
4.4. Pemodelan	48
4.5.1. Splitting.....	48
4.5.2. TF-IDF	49
4.5.3. SMOTE	50
4.5.4. Support Vector Machine	51
4.5.5. Naïve Bayes	54
4.5.6. Evaluasi Model	55
4.5. Visualisasi	59
4.6. Analisa Topic Modeling (LDA).....	63
4.7. Implementasi Hasil Model	66
BAB V PENUTUP	71
5.1 Kesimpulan.....	71
5.2 Saran.....	72

REFERENSI	73
LAMPIRAN.....	78
Link Deploy	78



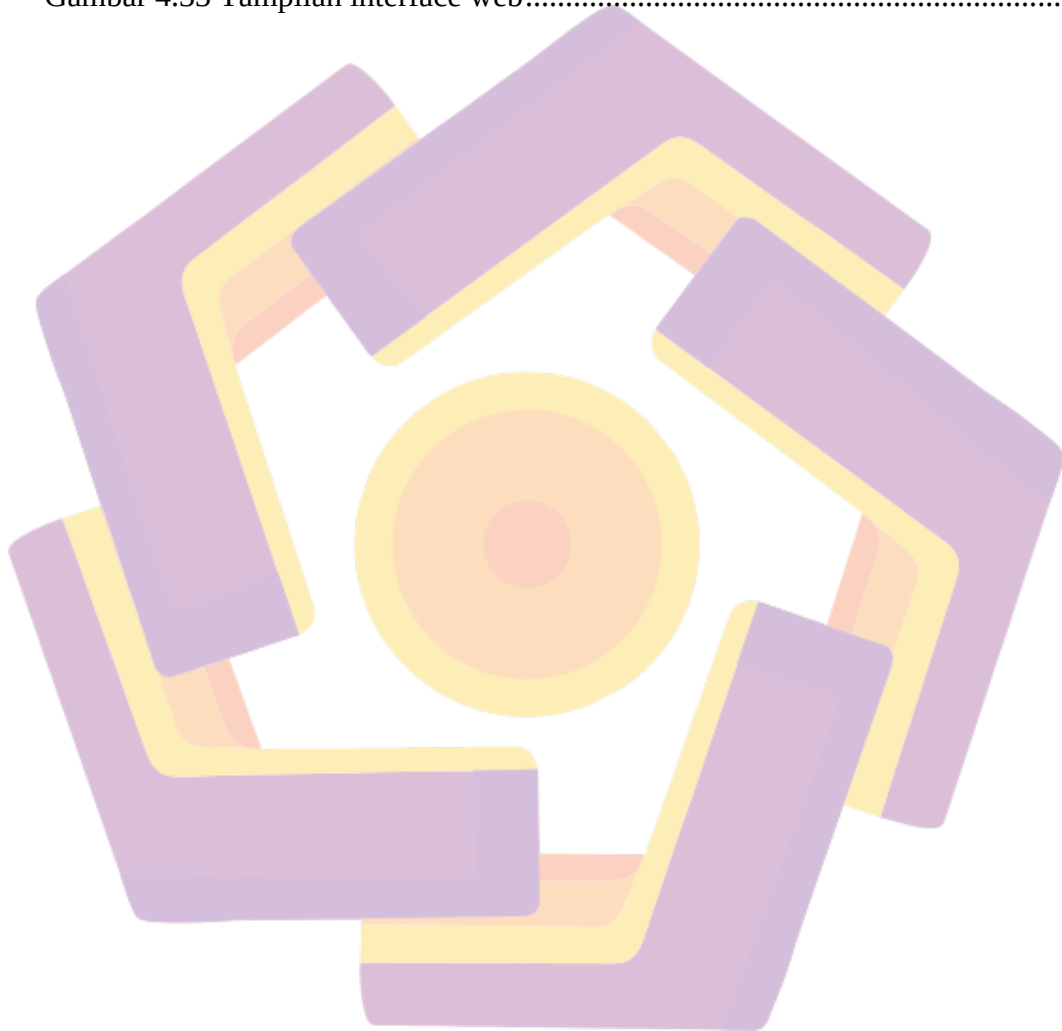
DAFTAR TABEL

Tabel 2.1 Keaslian Penelitian	9
Tabel 2.2 Confusion Matrix	18
Tabel 3.1 Struktur Dataframe	23
Tabel 3.2 Contoh tahap Cleaning.....	24
Tabel 3.3 Contoh tahap Case Folding.....	24
Tabel 3.4 Contoh tahap Normalisasi.....	25
Tabel 3.5 Contoh tahap Tokenize	25
Tabel 3.6 Contoh tahap Stopword Removal	26
Tabel 3.7 Contoh tahap Stemming.....	26
Tabel 3.8 Contoh teks Labeling	28
Tabel 3.9 Contoh Dataset.....	34
Tabel 4.1 Dataset Final	38
Tabel 4.2 Hasil Cleaning.....	40
Tabel 4.3 Hasil Case Folding.....	41
Tabel 4.4 Hasil Normalisasi.....	43
Tabel 4.5 Hasil Tokenize	44
Tabel 4.6 Hasil Stopword Removal	44
Tabel 4.7 Hasil Stemming.....	45
Tabel 4.8 Hasil Labeling.....	48
Tabel 4.9 Akurasi Support Vector Machine pada pembagian tertentu	53
Tabel 4.10 Waktu eksekusi Support Vector Machine pada pembagian tertentu ...	53
Tabel 4.11 Akurasi Naïve Bayes pada pembagian tertentu	55
Tabel 4.12 Waktu eksekusi Naïve Bayes pada pembagian tertentu	55
Tabel 4.13 Confusion matrix Support Vector Machine.....	56
Tabel 4.14 Confusion matrix Naïve Bayes	57
Tabel 4.15 Tabel distribusi sentimen	61
Tabel 4.16 Tabel hasil LDA.....	64

DAFTAR GAMBAR

Gambar 2.1 Pemisahan dua kelas (-1 dan +1)	16
Gambar 3.1 Alur penelitian.....	22
Gambar 4.1 Script crawling	37
Gambar 4.2 Script proses cleaning bagian 1	39
Gambar 4.3 Script proses cleaning bagian 2	40
Gambar 4.4 Script proses case folding	41
Gambar 4.5 Script proses normalisasi bagian 1	42
Gambar 4.6 Script proses normalisasi bagian 2	42
Gambar 4.7 Script proses tokenizing	43
Gambar 4.8 Script proses stopword removal	44
Gambar 4.9 Script proses stemming	45
Gambar 4.10 Script proses data quality enhancement	46
Gambar 4.11 Script proses labeling	47
Gambar 4.12 Wordcloud hasil labeling	48
Gambar 4.13 Script proses Splitting	48
Gambar 4.14 Script proses TF-IDF	49
Gambar 4.15 Script proses SMOTE	50
Gambar 4.16 Hasil proses SMOTE	51
Gambar 4.17 Script proses Support Vector Machine	52
Gambar 4.18 Script proses Naïve Bayes	54
Gambar 4.19 Script proses evaluasi model Support Vector Machine	56
Gambar 4.20 Classification report Support Vector Machine	56
Gambar 4.21 Script proses evaluasi model Naïve Bayes	57
Gambar 4.22 Classification report Naïve Bayes	57
Gambar 4.23 Grafik perbandingan akurasi model	59
Gambar 4.24 Grafik perbandingan waktu eksekusi model	60
Gambar 4.25 Grafik distribusi sentimen tiap bulan	61
Gambar 4.26 Pie Chart Sentimen	62
Gambar 4.27 Script proses LDA bagian 1	63

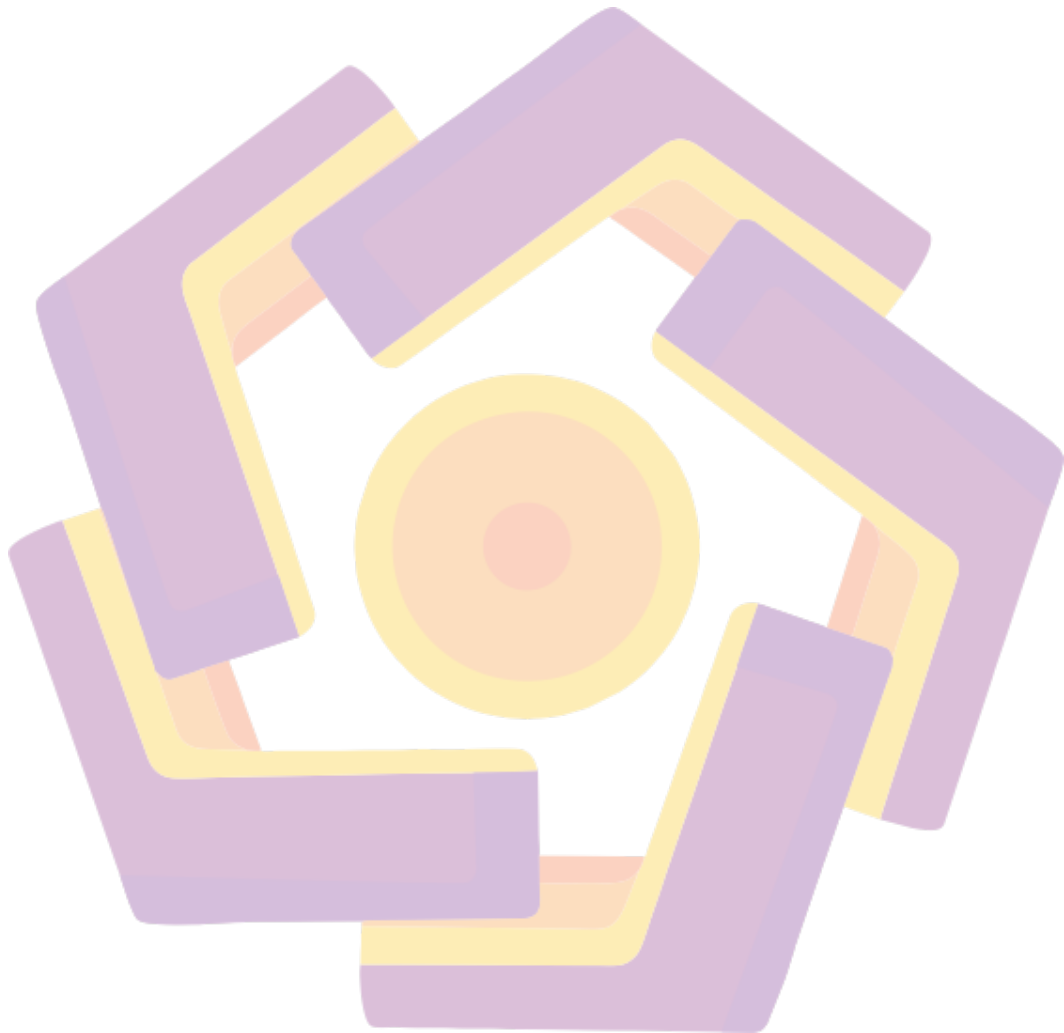
Gambar 4.28 Script proses LDA bagian 2	64
Gambar 4.29 Script proses menyimpan model	66
Gambar 4.30 Struktur folder	67
Gambar 4.31 Script full code app.py	68
Gambar 4.32 Script full code index.html	69
Gambar 4.33 Tampilan interface web.....	70



DAFTAR LAMPIRAN

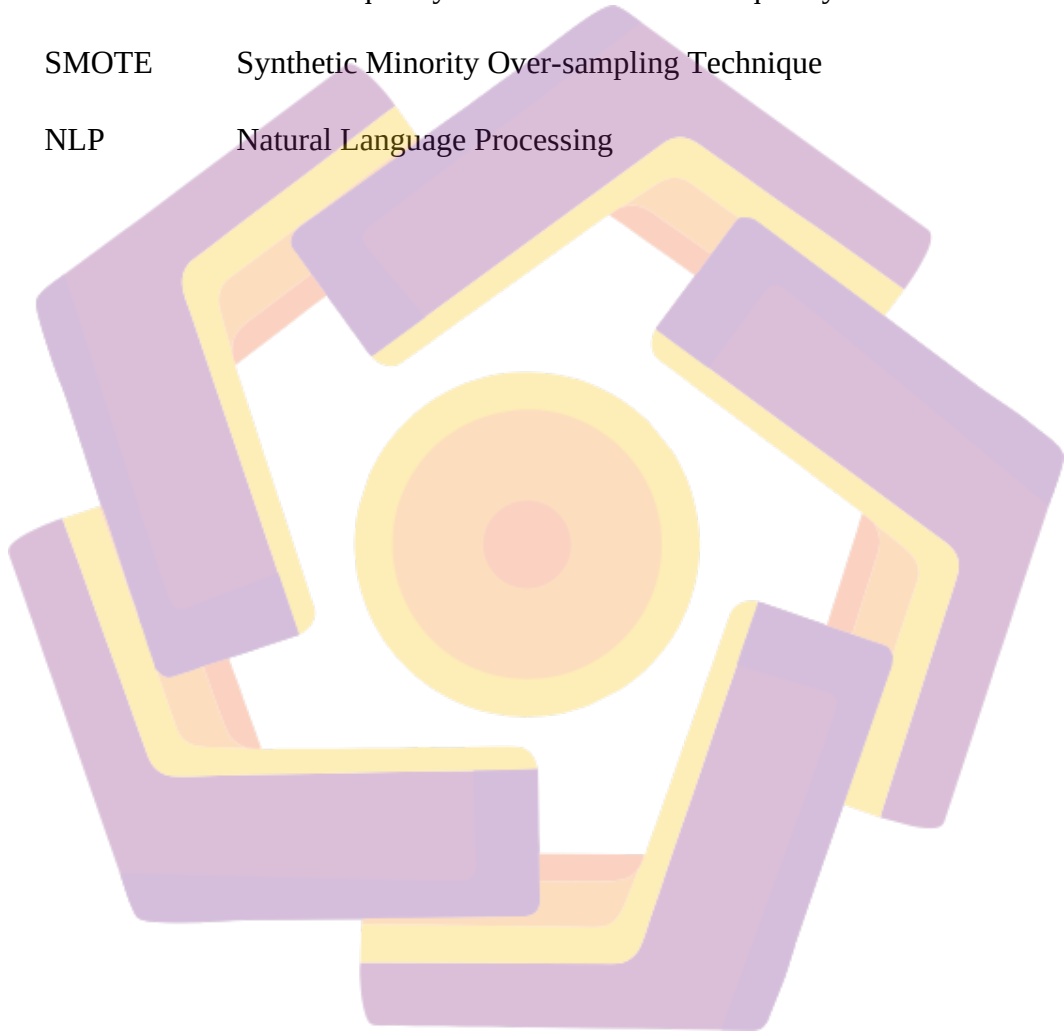
Lampiran 1. Link Deploy

78



DAFTAR LAMBANG DAN SINGKATAN

SVM	Support Vector Machines
LDA	Latent Dirichlet Allocation
TF-IDF	Term Frequency-Inverse Document Frequency
SMOTE	Synthetic Minority Over-sampling Technique
NLP	Natural Language Processing



INTISARI

Masa transisi pemerintahan sering diwarnai dengan beragam opini publik yang terbagi, terutama di platform media sosial seperti YouTube. Studi ini menganalisis komentar dari dua belas video yang berkaitan dengan masa transisi pemerintahan Presiden Joko Widodo untuk memahami sentimen publik. Data yang terkumpul melalui komentar tersebut diproses melalui tahapan pembersihan, tokenisasi, normalisasi, dan tahapan lainnya untuk menghasilkan data yang siap dianalisis.

Metode Term Frequency-Inverse Document Frequency (TF-IDF) dilakukan untuk pembobotan teks dan data diseimbangkan menggunakan Synthetic Minority Over-sampling Technique (SMOTE). Algoritma Support Vector Machine (SVM) dan Naïve Bayes digunakan untuk melakukan klasifikasi sentimen pada penelitian ini. Hasil penelitian menunjukkan bahwa SVM menghasilkan akurasi yang lebih tinggi (97,60%) dibandingkan Naïve Bayes (90,54%), meskipun SVM memerlukan waktu eksekusi yang lebih lama. Selain itu, beberapa topik penting dalam opini publik hasil penelitian seperti ketidakpuasan terhadap kebijakan, nepotisme, kebijakan pembangunan, dan harapan terhadap pemerintahan baru diidentifikasi menggunakan Latent Dirichlet Allocation (LDA).

Hasil penelitian ini diharapkan dapat membantu pembuat kebijakan memahami lebih baik mengenai persepsi publik. Selain itu, temuan ini juga dapat digunakan sebagai referensi akademis untuk mengembangkan algoritma analisis sentimen yang lebih lanjut. Penelitian berikutnya disarankan untuk menggunakan dataset yang lebih luas dan algoritma modern seperti DeBERTa atau XLNet untuk meningkatkan akurasi dan efisiensi.

Kata kunci: Analisis Sentimen, Masa Transisi, Support Vector Machine, Naïve Bayes, YouTube.

ABSTRACT

The transition of government is often accompanied by divided public opinion, especially on social media platforms like YouTube. This study analyzes comments from twelve videos related to the transition period of President Joko Widodo's administration to understand public sentiment. The data collected from these comments were processed through stages of cleaning, tokenization, normalization, and other steps to prepare the data for analysis.

The Term Frequency-Inverse Document Frequency (TF-IDF) method was employed for text weighting, and data balancing was performed using the Synthetic Minority Over-sampling Technique (SMOTE). The Support Vector Machine (SVM) and Naïve Bayes algorithms were used for sentiment classification in this study. The results show that SVM achieved higher accuracy (97.60%) compared to Naïve Bayes (90.54%), although SVM required more execution time. Furthermore, several key topics in public opinion, such as dissatisfaction with policies, nepotism, development policies, and expectations of the new government, were identified using Latent Dirichlet Allocation (LDA).

This study's findings are expected to help policymakers better understand public perceptions. Additionally, these findings can serve as an academic reference for further development of sentiment analysis algorithms. Future research is recommended to use a larger dataset and modern algorithms such as DeBERTa or XLNet to improve accuracy and efficiency.

Keyword: Sentiment Analysis, Government Transition, Support Vector Machine, Naïve Bayes, YouTube.