

**ANALISIS SENTIMEN TERHADAP GIBRAN SEBAGAI CALON
WAKIL PRESIDEN MENGGUNAKAN ALGORITMA SUPPORT
VECTOR MACHINE (SVM) DAN ADABOOST DENGAN
METODE PENDEKATAN ZERO-SHOT**

SKRIPSI

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



disusun oleh

DZULFIQAR MUHAMMAD

20.61.0184

Kepada

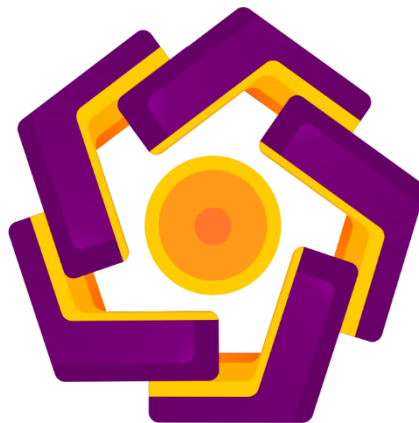
**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2024

**ANALISIS SENTIMEN TERHADAP GIBRAN SEBAGAI CALON
WAKIL PRESIDEN MENGGUNAKAN ALGORITMA SUPPORT
VECTOR MACHINE (SVM) DAN ADABOOST DENGAN
METODE PENDEKATAN ZERO-SHOT**

SKRIPSI

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



disusun oleh

DZULFIQAR MUHAMMAD

20.61.0184

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2024

HALAMAN PERSETUJUAN

SKRIPSI

**ANALISIS SENTIMEN TERHADAP GIBRAN SEBAGAI CALON WAKIL
PRESIDEN MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE
(SVM) DAN ADABOOST DENGAN METODE PENDEKATAN ZERO-SHOT**

yang disusun dan diajukan oleh

Dzulfiqar Muhammad

20.61.0184

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal Kamis, 22 Agustus 2024

Dosen Pembimbing,



Hastari Utama, M.Cs
NIK. 190302230

HALAMAN PENGESAHAN

SKRIPSI

ANALISIS SENTIMEN TERHADAP GIBRAN SEBAGAI CALON WAKIL PRESIDEN MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE (SVM) DAN ADABOOST DENGAN METODE PENDEKATAN ZERO-SHOT

yang disusun dan diajukan oleh

Dzulfiqar Muhammad

20.61.0184

Telah dipertahankan di depan Dewan Penguji
pada tanggal Kamis, 22 Agustus 2024

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Sri Ngudi Wahyuni, S.T., M.Kom., Dr.
NIK. 190302060



Theopilus Bayu Sasongko, S.Kom, M.Eng
NIK. 190302375



Hastari Utama, M.Cs
NIK. 190302230



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal Kamis, 22 Agustus 2024

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 : **Dzulfiqar Muhammad**
NIM : **20.61.0184**

Menyatakan bahwa Skripsi dengan judul berikut:

**ANALISIS SENTIMEN TERHADAP GIBRAN SEBAGAI CALON WAKIL
PRESIDEN MENGGUNAKAN ALGORITMA SUPPORT VECTOR
MACHINE (SVM) DAN ADABOOST DENGAN METODE PENDEKATAN
ZERO-SHOT**

Dosen Pembimbing : **Hastari Utama, M.Cs**

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, Kamis, 22 Agustus 2024

Yang Menyatakan



94ALX038002726
D. **Dzulfiqar Muhammad**

KATA PENGANTAR

Puji syukur saya ucapkan kepada Allah SWT karena atas nikmat dan rahmat Nya, penulis dapat menyelesaikan skripsi ini sebagai syarat untuk memperoleh gelar Sarjana Komputer Program Studi Informatika Fakultas Ilmu Komputer Universitas Amikom Yogyakarta Selama proses penyusunan skripsi ini, tidak terlepas dari beberapa bantuan, saran, kritik dan dukungan. Dalam kesempatan ini, penulis berterima kasih kepada:

1. Kepada Bapak Hastari Utama, M.Cs selaku dosen pembimbing yang telah meluangkan waktu, memberikan bantuan, bimbingan, dan motivasi.
2. Kedua orang tua penulis Bapak M. Damar Yowono, S.Pd dan Ibu Sri Rahayu yang selalu membimbing, mendoakan dan mendukung dalam setiap langkah yang akan dilalui.
3. Seluruh sahabat dan teman yang saling memberikan semangat dan partisipasinya selama proses menuju sidang.
4. Seluruh pihak yang secara langsung maupun tidak langsung telah mendukung dan mendoakan saya dalam pengerjaan skripsi ini.

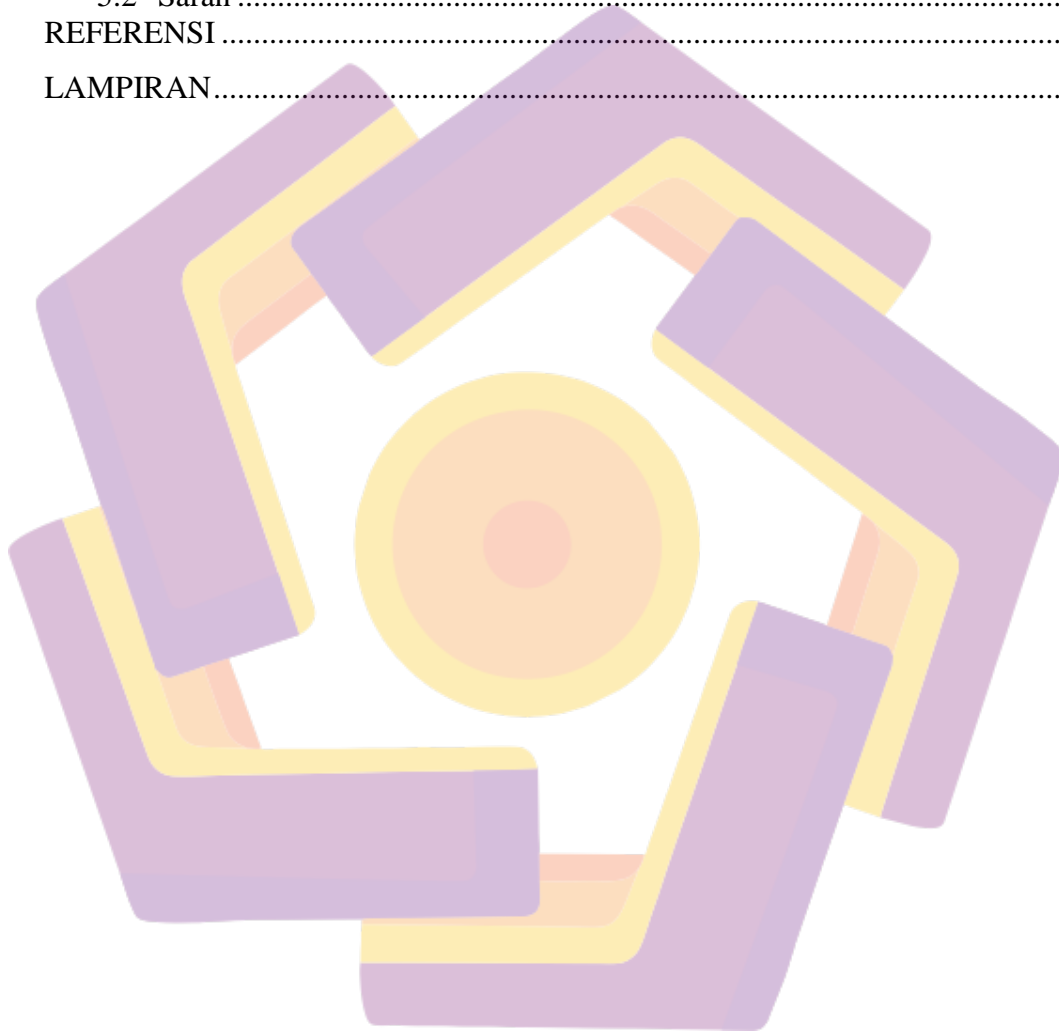
Yogyakarta, Kamis, 22 Agustus 2024

Dzulfiqar Muhammad

DAFTAR ISI

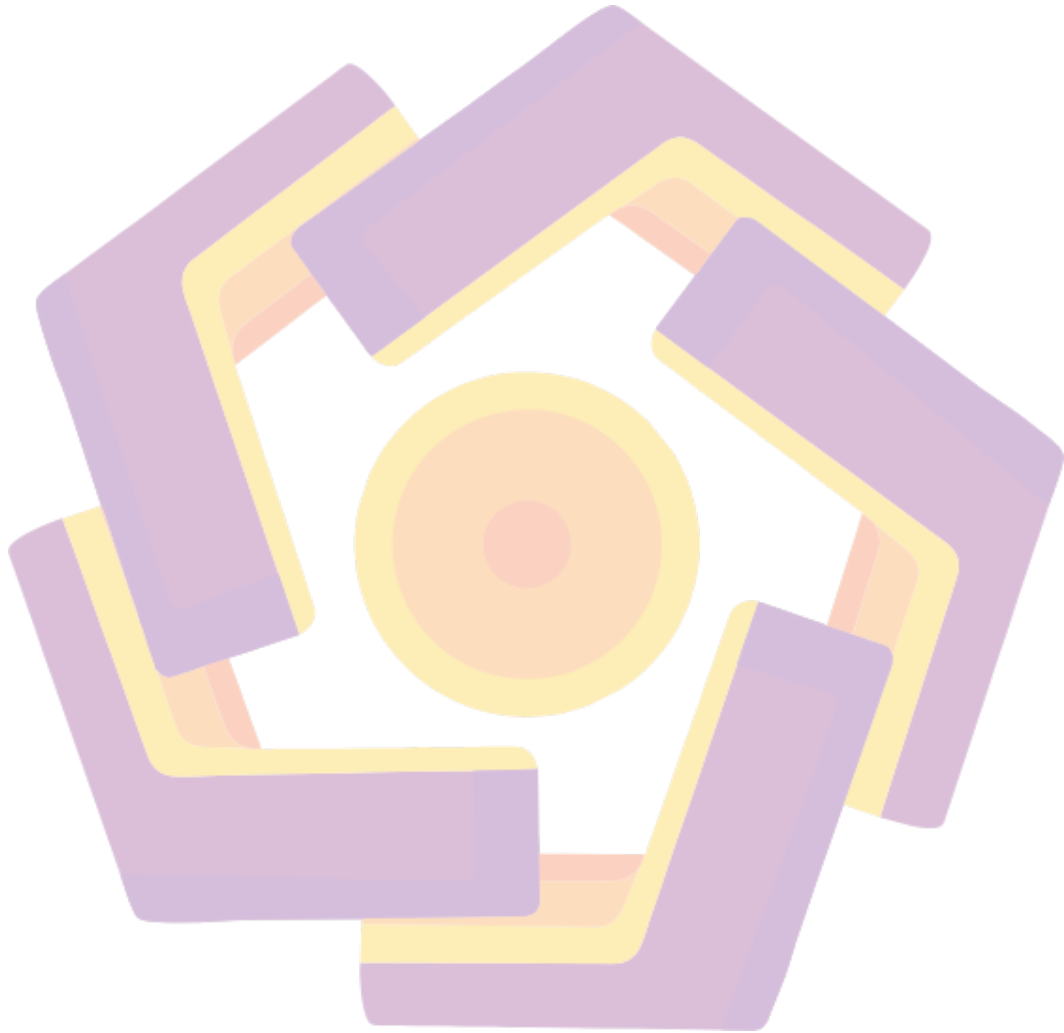
HALAMAN JUDUL	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
KATA PENGANTAR	v
DAFTAR ISI.....	vi
DAFTAR TABEL.....	viii
DAFTAR GAMBAR	ix
DAFTAR LAMPIRAN.....	x
DAFTAR LAMBANG DAN SINGKATAN	xi
DAFTAR ISTILAH	xii
INTISARI	xiii
ABSTRACT.....	xiv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	3
1.3 Batasan Masalah	4
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	4
1.6 Sistematika Penulisan	5
BAB II TINJAUAN PUSTAKA	6
2.1 Studi Literatur	6
2.2 Dasar Teori	17
BAB III METODE PENELITIAN	30
3.1 Objek Penelitian.....	30
3.2 Alur Penelitian	30
3.3 Alat dan Bahan.....	38
BAB IV HASIL DAN PEMBAHASAN	39
4.1 Persiapan Google Colab	39
4.2 <i>Harvesting Data tweets</i>	39
4.2 <i>Data Pre-processing</i>	42
4.4 Model Zero-Shot Classification	44
4.5 Menyaring Data Dengan Skor lebih dari 70%	45
4.6 Eksplorasi Data Analisis (EDA).....	46

4.7 Pembobotan Kata Dengan TF-IDF.....	49
4.8 Perhitungan Akurasi Model Dengan Kernel SVM.....	51
4.9 Metrik Evaluasi dengan <i>Cross-validate</i>	54
4.10 Kolaborasi SVM dengan <i>AdaBoost</i> Sebagai <i>Ensemble Learning</i>	55
4.11 Menerapkan Model Kedalam Data Baru.....	56
BAB V PENUTUP	60
5.1 Kesimpulan	60
5.2 Saran	60
REFERENSI	62
LAMPIRAN.....	66



DAFTAR TABEL

Tabel 2. 1 Perbandingan Metode	16
Tabel 4. 1 Hasil dan rata rata implementasi kernel SVM.....	53



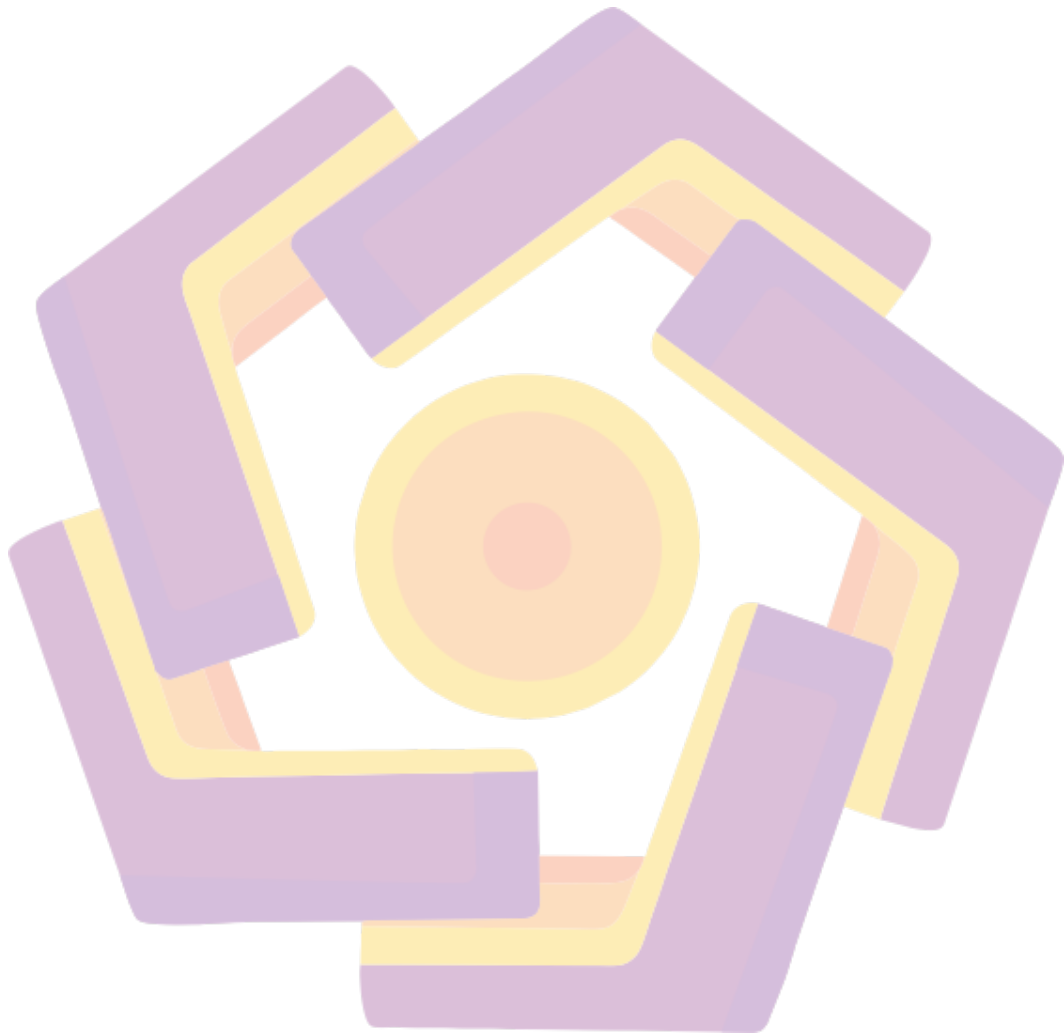
DAFTAR GAMBAR

Gambar 2. 1 Ilustrasi SVM	23
Gambar 2. 2 Cara kerja SVM.....	24
Gambar 2. 3 Ilustrasi Adaboost	25
Gambar 2. 4 Demo model <i>Zero-Shot</i>	28
Gambar 3. 1 Alur penelitian.....	31
Gambar 4. 1 Tampilan awal <i>google colab</i>	39
Gambar 4. 2 Hasil <i>harvesting data</i>	41
Gambar 4. 3 Proses pembersihan data	42
Gambar 4. 4 Program dan hasil hapus data duplikat dan kosong	43
Gambar 4. 5 Jumlah data	43
Gambar 4. 6 Hasil preprocessing data	43
Gambar 4. 7 Program pemberian label	44
Gambar 4. 8 Hasil pemberian skor pada teks.....	45
Gambar 4. 9 Program <i>Filtering Data</i>	45
Gambar 4. 10 Hasil Filtering	46
Gambar 4. 11 Diagram batang hasil tiap labelnya.....	47
Gambar 4. 12 Contoh hasil sentimen positif.....	48
Gambar 4. 13 Contoh hasil sentimen netral.....	48
Gambar 4. 14 Contoh hasil sentimen negatif.....	49
Gambar 4. 15 Program pembobotan kata	49
Gambar 4. 16 Hasil pembobotan kata.....	50
Gambar 4. 17 Kata dengan bobot tertinggi	50
Gambar 4. 18 Program perhitungan akurasi	51
Gambar 4. 19 Implementasi kernel SVM polynomial	52
Gambar 4. 20 Hasil implementasi kernel SVM.....	52
Gambar 4. 21 Program kolaborasi SVM dengan Adaboost.....	55
Gambar 4. 22 Hasil kolaborasi SVM dengan Adaboost.....	56
Gambar 4. 23 Hasil preprocessing data baru	57
Gambar 4. 24 Program penerapan model kedalam data baru	57
Gambar 4. 25 Hasil penerapan data baru	58
Gambar 4. 26 Perbandingan jumlah hasil data baru	58
Gambar 4. 27 Pie chart perbandingan hasil data baru.....	59
Gambar 5. 1 Dataset penelitian	66

DAFTAR LAMPIRAN

Lampiran 1. Dataset penelitian

67

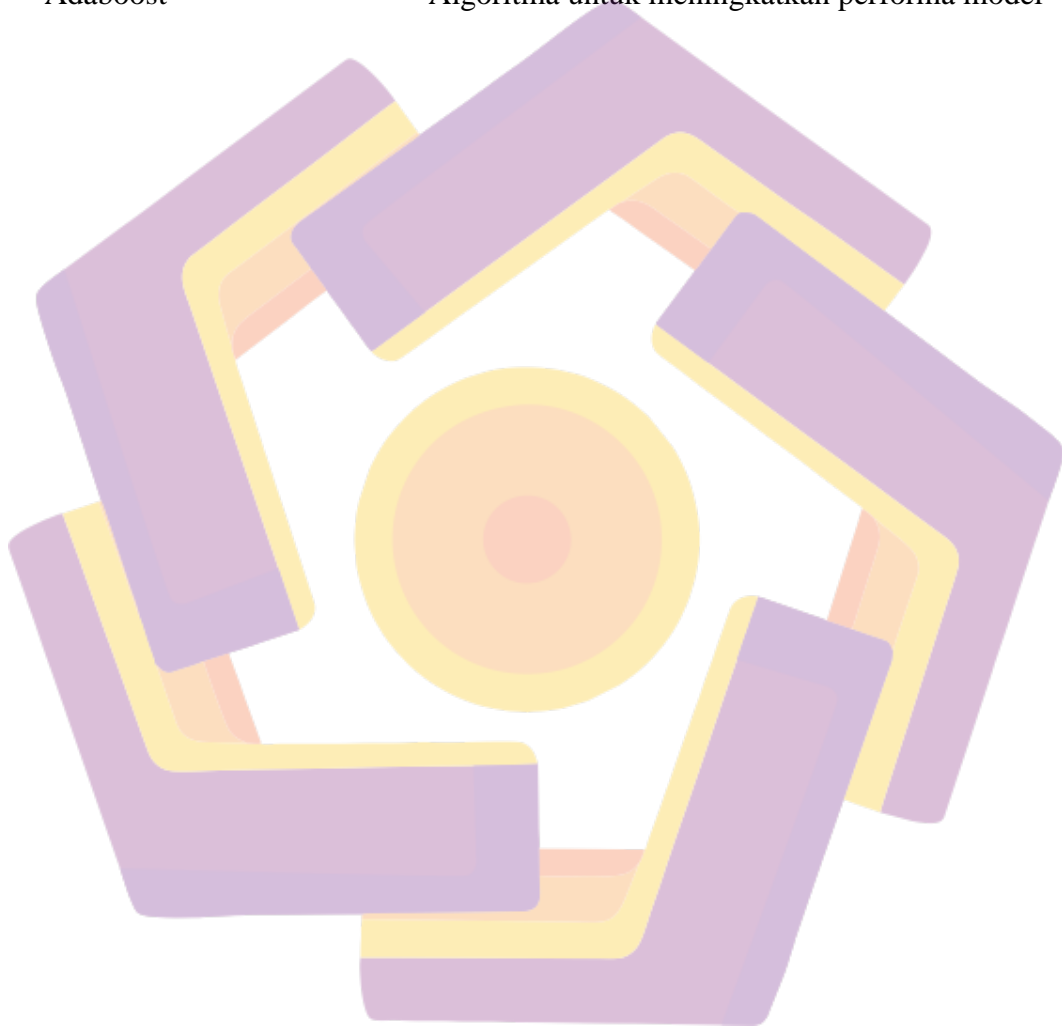


DAFTAR LAMBANG DAN SINGKATAN

SVM	Support Vector Machines
BERT	Bidirectional Encoder Representations from Transformers
TF-IDF	Term Frequency-Inverse Document Frequency
NLP	Natural Language Processing
MK	Mahkamah Konstitusi
AI	Artificial Intelligence
MLM	Masked Language Model
NSP	Next Sentence Prediction
AdaBoost	Adaptive Boosting
CSV	Comma Separated Values
RE	Regular Expression
RBF	Radial Basis Function

DAFTAR ISTILAH

Analisis Sentimen	Teknik untuk mengidentifikasi opini dalam teks
Support Vector Machine	Algoritma untuk klasifikasi
Zero-shot	Model untuk mengklasifikasikan teks
Adaboost	Algoritma untuk meningkatkan performa model



INTISARI

Penelitian ini bertujuan untuk menganalisis sentimen publik terhadap pencalonan Gibran Rakabuming Raka sebagai calon Wakil Presiden, yang menuai berbagai perdebatan di masyarakat Indonesia, melalui data teks yang diperoleh dari *tweet* dari media sosial *Twitter* atau yang saat ini disebut dengan *X*. Analisis sentimen ini bertujuan untuk mengidentifikasi atau memberi label apakah teks yang dianalisis mengandung sentimen positif, negatif, atau netral. Pelabelan manual dengan data yang besar oleh manusia rentan terhadap inkonsistensi karena berbagai faktor seperti kelelahan dan perbedaan pemahaman. Pemberian label dengan menggunakan model *Zero-Shot* dapat meningkatkan efisiensi dan kualitas hasil analisis. Evaluasi model dalam analisis sentimen dilakukan secara otomatis dengan menggunakan algoritma *Support Vector Machine* (SVM) dan metode *ensemble learning AdaBoost*. Dari total keseluruhan 27.000 data *tweet*, proses *pre-processing* menghasilkan 22.915 *tweet* yang akan dianalisis dengan model *Zero-Shot*. Dengan menetapkan treshold 70%, data disaring menjadi 13.932 *tweet* untuk meningkatkan tingkat keakuratan analisis. Evaluasi kinerja model SVM dengan berbagai kernel menunjukkan bahwa kernel linear memberikan hasil terbaik dengan rata-rata akurasi 91,42%. Hasil akurasi sebesar 91,42%. *Precision* dan *recall*, masing-masing 91,49% dan 91,42%, nilai *F1-score* lebih rendah yaitu sebesar 90,99%. Kolaborasi antara SVM kernel linear dengan algoritma *ensemble Adaboost* menghasilkan rata-rata akurasi 0.6828. Penelitian ini juga melakukan penerapan model kedalam data baru sebanyak 2.777 *tweet* menunjukkan dominasi sentimen negatif dengan 2.324 *tweet*, sementara hanya 453 *tweet* yang tergolong positif.

Kata kunci: *tweet, Twitter, X, Zero-shot, Support Vector Machine, AdaBoost*

ABSTRACT

This research aims to analyze public sentiment towards the candidacy of Gibran Rakabuming Raka as a Vice Presidential candidate, which has sparked various debates within Indonesian society, through text data obtained from tweets on social media platform Twitter, now known as X. The sentiment analysis seeks to identify and label whether the analyzed texts contain positive, negative, or neutral sentiments. Manual labeling of large datasets by humans is prone to inconsistencies due to factors such as fatigue and differing interpretations. Using a Zero-Shot model for labeling can enhance the efficiency and quality of the analysis results. The evaluation of the model in sentiment analysis is conducted automatically using a Support Vector Machine (SVM) algorithm and the ensemble learning method AdaBoost. Out of a total of 27,000 tweets, the pre-processing stage resulted in 22,915 tweets that were analyzed using the Zero-Shot model. By setting a threshold of 70%, the data was filtered down to 13,932 tweets to improve the accuracy of the analysis. The evaluation of the SVM model's performance with various kernels showed that the linear kernel provided the best results with an average accuracy of 0.9142. The accuracy result is 91.42% precision and recall are 91.49% and 91.42%, respectively, while the F1-score is lower at 90.99%. The collaboration between the linear SVM kernel and AdaBoost resulted in an average accuracy of 0.6828. The application of the model to a new dataset of 2,777 tweets revealed a dominance of negative sentiment with 2,324 tweets, while only 453 tweets were classified as positive.

Keyword: *tweet, Twitter, X, Zero-shot, Support Vector Machine, AdaBoost*