

**ANALISIS SENTIMEN KOMENTAR PENGGUNA  
INSTAGRAM MENGENAI PELAKSANAAN PEMILU  
2024 DENGAN NAIVE BAYES DAN LEXICON-BASED**

**SKRIPSI SARJANA SISTEM INFORMASI**

Oleh

Cahyani Rahma Dewi

207006516092



**PROGRAM STUDI SISTEM INFORMASI  
FAKULTAS TEKNOLOGI KOMUNIKASI DAN  
INFORMATIKA**

**UNIVERSITAS NASIONAL**

**2023**

**Analisis Sentimen Komentar Pengguna Instagram  
mengenai Pelaksanaan Pemilu 2024 dengan Naive Bayes  
dan Lexicon-Based**

Karya ilmiah sebagai salah satu syarat untuk memperoleh gelar  
Sarjana Teknik Informatika dari Fakultas Teknologi Komunikasi dan Informatika

Oleh

Cahyani Rahma Dewi

207006516092



**PROGRAM STUDI SISTEM INFORMASI**  
**FAKULTAS TEKNOLOGI KOMUNIKASI DAN**  
**INFORMATIKA**  
**UNIVERSITAS NASIONAL**

**2023**

**HALAMAN PENGESAHAN**

HALAMAN PENGESAHAN

TUGAS AKHIR

**Analisis Sentimen Komentar Pengguna Instagram mengenai  
Pelaksanaan Pemilu 2024 dengan Naïve Bayes dan Lexicon-Based**



Cahyani Rahma Dewi

207006516092

UNIVERSITAS NASIONAL

Dosen Pembimbing I

(Agus Iskandar, S.Kom., M.Kom)

## HALAMAN KEASLIAN TUGAS AKHIR

### PERNYATAAN KEASLIAN TUGAS AKHIR

Saya menyatakan dengan sesungguhnya bahwa Tugas Akhir dengan judul :

**Analisis Sentimen Komentar Pengguna Instagram mengenai  
Pelaksanaan Pemilu 2024 dengan Naïve Bayes dan Lexicon-Based**

Yang dibuat untuk melengkapi salah satu persyaratan menjadi Sarjana Komputer pada Program Studi Sistem Informasi, Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional, sebagaimana yang saya ketahui adalah bukan merupakan tiruan atau publikasi dari Tugas Akhir yang pernah diajukan atau dipakai untuk mendapatkan gelar di lingkungan Universitas Nasional maupun perguruan tinggi atau instansi lainnya, kecuali pada bagian – bagian tertentu yang menjadi sumber Informasi atau acuan yang dicantumkan sebagaimana mestinya.

Jakarta, 26 Agustus 2024



Cahyani Rahma Dewi

NPM : 207006516092

## LEMBAR PERSETUJUAN TUGAS AKHIR

### LEMBAR PERSETUJUAN TUGAS AKHIR

Tugas Akhir dengan judul :

**Analisis Sentimen Komentar Pengguna Instagram mengenai Pelaksanaan Pemilu 2024 dengan Naïve Bayes dan Lexicon-Based**

Dibuat untuk melengkapi salah satu persyaratan menjadi Sarjana Komputer pada Program Studi Sistem Informasi, Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional. Tugas Akhir ini diujikan pada Sidang Akhir Semester Genap 2023-2024 pada tanggal 22 Agustus 2024

**Dosen Pembimbing 1**

  
Agus Iskandar, S.Kom., M.Kom  
NIDN. 0310087503

**Ketua Program Studi**

  
Dr. Andrianingsih, S.Kom., MMSI  
NIDN. 030309790



**LEMBAR PERSETUJUAN JUDUL PEMBIMBING 1**  
**LEMBAR PERSETUJUAN JUDUL YANG TIDAK ATAU YANG DIREVISI**

Nama : Cahyani Rahma Dewi

NPM : 207006516092

Fakultas/Akademi : Fakultas Teknologi Komunikasi dan Informatika

Program Studi : Sistem Informasi

Tanggal Sidang : 22 Agustus 2024

JUDUL DALAM BAHASA INDONESIA :

**Analisis Sentimen Komentar Pengguna Instagram mengenai Pelaksanaan Pemilu 2024 dengan Naive Bayes dan Lexicon-Based**

JUDUL DALAM BAHASA INGGRIS :

**Sentiment Analysis of Instagram User Comments on the Implementation of the 2024 Election using Naive Bayes and Lexicon-Based**

**TANDA TANGAN DAN TANGGAL**

Pembimbing 1	Ka. Prodi	Mahasiswa
TGL:	TGL: 26.08.2024	TGL:
	 UNIVERSITAS NASIONAL	 Cahyani Rahma Dewi

## KATA PENGANTAR

Puji dan syukur penulis panjatkan kepada Tuhan Yang Maha Esa atas segala rahmat dan karunia-Nya yang telah memberikan kemudahan serta petunjuk dalam perjalanan penulisan skripsi ini. Skripsi berjudul "**Analisis Sentimen Komentar Pengguna Instagram mengenai Pemilu 2024 dengan Naive Bayes dan Lexicon-Based**" telah berhasil diselesaikan sebagai bagian dari syarat kelulusan Program Studi Sarjana di Fakultas Teknologi Komunikasi dan Informatika.

Pelaksanaan riset dan penulisan tesis ini merupakan hasil dari kerja sama berbagai individu. Oleh karena itu, penulis ingin menyampaikan ucapan terima kasih yang besar, terutama kepada dosen pembimbing Tugas Akhir, yaitu Ibu Panca Dewi Pamungkasari, S.T.,M.T. Ph.D yang telah memberikan bantuan dan dukungan yang luar biasa. Mereka telah menghabiskan waktu, energi, memberikan pemikiran, memberikan panduan, petunjuk, semangat, serta pengertian terhadap segala keterbatasan yang penulis hadapi selama pelaksanaan riset tugas akhir dan penulisan skripsi. Penulis juga ingin menyampaikan rasa terima kasih yang tulus kepada:

1. Kepada Ibu Sugiyari dan Bapak Suyadi, orang yang terkasih dan tak pernah lelah memberikan dukungan, doa, restu, serta bantuan tanpa batas kepada penulis.
2. Seluruh dosen pengajar di Program Studi Sistem Informasi FTKI serta dosen dari program studi lain yang telah berbagi pengetahuan yang berharga.
3. Rekan seangkatan dan sesama mahasiswa dari berbagai angkatan yang telah memberikan dukungan yang berarti.
4. Kepada semua individu yang telah membantu dan memberikan bantuan yang sangat berarti dalam perjalanan penyusunan skripsi ini.

Penulis berharap Tuhan Yang Maha Esa membalas kebaikan dan bantuan yang diberikan dengan yang lebih baik. Kritik dan saran yang konstruktif sangat diharapkan, semoga skripsi ini memberikan manfaat di bidang Sistem Informasi.

Jakarta, 21 November 2023



Cahyani Rahma Dewi





## ABSTRAK

Perdebatan seputar pelaksanaan Pemilihan Umum 2024 yang menjadi sorotan utama di Indonesia, terutama dalam platform media sosial yang diminati oleh masyarakat. Perubahan pemimpin di Indonesia dan perbedaan emosional yang muncul dalam masyarakat menjadi perhatian penting. Pencarian figur kepemimpinan memunculkan berbagai perspektif teoritis, konseptual, dan budaya yang kompleks. Penulisan ini bertujuan menganalisis sentiment Masyarakat terkait dengan pelaksanaan pemilihan umum 2024 dengan mengklasifikasikan sentimen sebagai positif, negatif, atau netral, membantu pemahaman tentang persepsi masyarakat terhadap para kandidat, isu-isu politik yang relevan, serta pola perilaku pemilih. Metodologi yang digunakan melibatkan pengumpulan data dengan teknik *Scrapping* dari platform media sosial, yaitu Instagram. Dengan menggunakan kombinasi dari kedua algoritma, yaitu Naïve Bayes Classifier dan pelabelan Lexicon-Based metode ini digunakan untuk melakukan analisis sentimen terhadap pemilihan umum dalam penelitian ini. Penggabungan kedua pendekatan ini diharapkan memberikan pemahaman lebih mendalam tentang sentimen Masyarakat dalam komentar pengguna instagram.

**Kata Kunci:** Analisis Sentimen, Pemilihan Umum, Media Sosial, Opini Publik

## ABSTRACT

*The debate surrounding the implementation of the 2024 General Election has been the focal point in Indonesia, particularly on social media platforms favored by the public. The change in leadership in Indonesia and the emotional differences that have emerged within society have become significant concerns. The quest for leadership figures has brought forth various complex theoretical, conceptual, and cultural perspectives. This paper aims to analyze public sentiment regarding the conduct of the 2024 general election by classifying sentiment as positive, negative, or neutral, aiding in understanding public perceptions of the candidates, relevant political issues, and voter behavior patterns. The methodology involves data collection through crawling techniques from social media platforms, especially Twitter. By utilizing a combination of both algorithms, namely Naïve Bayes Classifier and Lexicon-Based labeling, this method is employed to conduct sentiment analysis on the general election in this study. The integration of these two approaches is expected to provide a deeper understanding of public sentiment on the Twitter social media platform with a dataset of 2000 entries. With this dataset, comprehensive and accurate analysis is made possible in facing the 2024 general election.*

**Keywords:** *Sentiment Analysis, General Elections, Social Media, Public Opinion.*

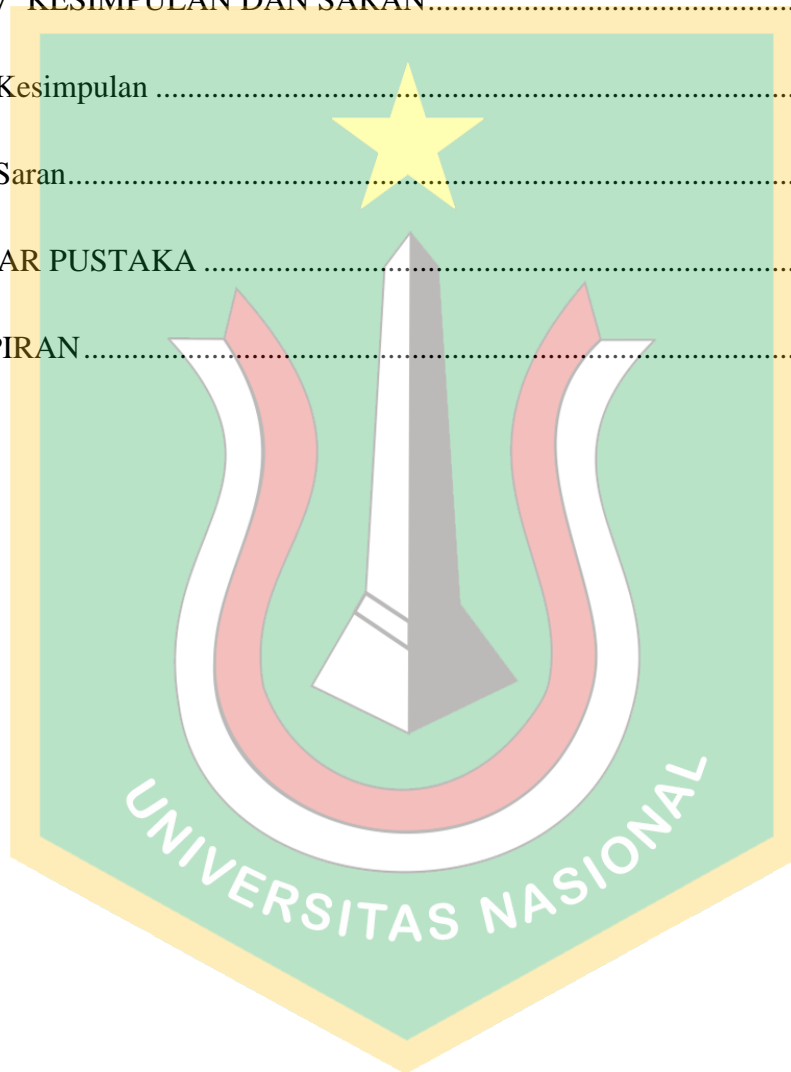


## DAFTAR ISI

KATA PENGANTAR .....	7
ABSTRAK .....	9
ABSTRACT .....	10
DAFTAR ISI .....	12
DAFTAR GAMBAR .....	15
DAFTAR TABEL .....	17
BAB I PENDAHULUAN .....	18
1.1 Latar Belakang .....	18
1.2 Identifikasi Masalah .....	20
1.3 Tujuan Penelitian .....	20
1.4 Batasan Masalah .....	21
1.5 Kontribusi .....	21
BAB II TINJAUAN PUSTAKA .....	22
2.1 Penelitian Terdahulu .....	22
2.2 Landasan Teori .....	32
2.2.1 Analisis Sentimen .....	32
2.2.3 Preprocessing .....	32
2.2.4 Pemilihan Fitur Lexicon Based .....	33

2.2.5	Pembobotan TF – IDF.....	34
2.2.6	Naïve Bayes .....	34
<b>BAB III METODE PENELITIAN.....</b>		<b>36</b>
3.1	Lokasi Penelitian.....	36
3.2	Penentuan Subjek Penelitian.....	36
3.3	Fokus penelitian .....	36
3.4	Sumber Data.....	37
3.5	Desain Penelitian.....	37
3.5.1	Scraping Data.....	40
3.5.2	Pre-Processing.....	40
3.5.3	Labelling .....	46
3.5.4	Pembobotan TF-IDF.....	47
3.5.5	Klasifikasi Naive Bayes.....	50
3.5.6	Evaluasi dan Result.....	54
<b>BAB IV HASIL DAN PEMBAHASAN.....</b>		<b>57</b>
4.1	Hasil .....	57
4.1.1	Web Scrapping.....	57
4.1.2	Pre-Processing.....	58
4.1.3	Pelabelan Lexicon Based .....	63
4.1.4	Ekstrasi Fitur TF-IDF.....	64

4.1.5	Klasifikasi Naïve Bayes .....	65
4.1.6	Evaluasi dan Result .....	66
4.2	Pembahasan .....	68
<b>BAB V KESIMPULAN DAN SARAN .....</b>		<b>73</b>
5.1	Kesimpulan .....	73
5.2	Saran .....	74
<b>DAFTAR PUSTAKA .....</b>		<b>75</b>
<b>LAMPIRAN .....</b>		<b>77</b>



## DAFTAR GAMBAR

Gambar 3 1 Alur Penelitian.....	38
Gambar 3 2 Data Scrapper .....	40
Gambar 3 3 Langkah Pre-Processing .....	41
Gambar 3 4 Contoh Word Cloud Positif .....	55
Gambar 3 5 Contoh Word Cloud Negatif .....	55
Gambar 3 6 Contoh Word Cloud Netral .....	55
Gambar 3 7 Contoh Bar Chart .....	56
Gambar 4. 1 Website Scrapper.....	57
Gambar 4. 2 Postingan Akun KPU .....	58
Gambar 4. 3 Data Komentari .....	58
Gambar 4. 4 Source Code Case Folding .....	59
Gambar 4. 5 Output Case Folding.....	60
Gambar 4. 6 Source Code Normalisasi.....	60
Gambar 4. 7 Output Normalisasi .....	61
Gambar 4. 8 Source Code Tokenizing .....	61
Gambar 4. 9 Output Tokenizing.....	61
Gambar 4. 10 Source Code Filtering .....	62
Gambar 4. 11 Output Filtering .....	62
Gambar 4. 12 Source Code Stemming.....	63
Gambar 4. 13 Output Stemming .....	63
Gambar 4. 14 Source Code Pelabelan Lexicon-Based.....	64
Gambar 4. 15 Output Pelabelan Lexicon Based .....	64

Gambar 4. 16 Source Code Tf-Idf .....	64
Gambar 4. 17 Output Tf-Idf .....	65
Gambar 4. 18 Source Code Naive Bayes .....	66
Gambar 4. 19 Source Code Confusion Matrix .....	66
Gambar 4. 20 Confusion Matrix .....	67
Gambar 4. 21 Sentiment Masyarakat .....	69
Gambar 4. 22 Output Accucary .....	70
Gambar 4. 23 Word Cloud Positif .....	71
Gambar 4. 24 Word Cloud Negatif .....	72
Gambar 4. 25 Word Cloud Netral .....	72





## DAFTAR TABEL

<b>Tabel 3.1</b> Casefolding.....	42
<b>Tabel 3.2</b> Normalization.....	42
<b>Tabel 3.3</b> Tokenization.....	43
<b>Tabel 3.4</b> Filtering (Stop word removal).....	44
<b>Tabel 3.5</b> Stemming .....	45
<b>Tabel 3.6</b> Kamus Lexicon Positive.....	46
<b>Tabel 3.7</b> Kamus Lexicon Negative.....	46
<b>Tabel 3.8</b> Hasil Labelling.....	47
<b>Tabel 3.9</b> Data Stemming.....	47
<b>Tabel 3.10</b> Proses perhitungan TF-IDF.....	48
<b>Tabel 3.11</b> Data tf-idf dan Sentimentnya.....	50

