



UNIVERSITAS  
**MERCU BUANA**

**ANALISIS DATA SENTIMEN TERKAIT OPINI HASIL  
PEMILU 2024 DI MEDIA SOSIAL DENGAN KOMPARASI  
MODEL KLASIFIKASI TEXT MINING**



UNIVERSITAS  
**MERCU BUANA**

Indah Fatimah Afriani

41820120034

**PROGRAM STUDI SISTEM INFORMASI  
FAKULTAS ILMU KOMPUTER UNIVERSITAS MERCU BUANA  
JAKARTA 2024**



UNIVERSITAS  
**MERCU BUANA**

**ANALISIS DATA SENTIMEN TERKAIT OPINI HASIL  
PEMILU 2024 DI MEDIA SOSIAL DENGAN KOMPARASI  
MODEL KLASIFIKASI TEXT MINING**

*Laporan Tugas Akhir*

Diajukan Untuk Melengkapi Salah Satu Syarat

Memperoleh Gelar Sarjana Komputer Oleh:

**Indah Fatimah Afriani**

**41820120034**

**PROGRAM STUDI SISTEM INFORMASI**

**FAKULTAS ILMU KOMPUTER**

**UNIVERSITAS MERCU BUANA**

**JAKARTA**

**2024**

## HALAMAN PERNYATAAN KARYA SENDIRI

Saya yang bertanda tangan di bawah ini:

Nama : Indah Fatimah Afriani  
NIM : 41820120034  
Pembimbing TA : Riri Fajriah, S.Kom, MM  
Judul : Analisis Data Sentimen Terkait Opini Hasil Pemilu 2024 di Media Sosial dengan Komparasi Model Klasifikasi Text Mining

Menyatakan bahwa Laporan Aplikatif/Tugas Akhir/Jurnal/Media Ilmiah ini adalah hasil karya saya sendiri dan bukan plagiat, serta semua sumber baik yang dikutip maupun yang dirujuk telah saya nyatakan dengan benar. Apabila ternyata ditemukan di dalam Laporan Tugas Akhir saya terdapat unsur plagiat, maka saya siap mendapatkan sanksi akademis yang berlaku di Universitas Mercu Buana.

Jakarta, 18 Januari 2024

  
Indah Fatimah Afriani

## LEMBAR PENGESAHAN

Nama Mahasiswa (1) : Indah Fatimah Afriani  
NIM (41820120034)  
Judul Tugas Akhir : Analisis Data Sentimen Terkait Opini Hasil Pemilu  
2024 di Media Sosial dengan Komparasi Model  
Klasifikasi Text Mining

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 18 Januari 2024

Menyetujui,

Pembimbing : Riri Fajriah S.Kom, MM.,   
NIDN : 0321108502  
Ketua Penguji : Dr. Bambang Jokonowo, S.Si, M.T.I. ()  
NIDN : 0320037002  
Penguji 1 : Dr. Bambang Jokonowo, S.Si, M.T.I. ()  
NIDN : 0320037002  
Penguji 2 : Inna Sabily Karima, S.Kom, M.Kom ()  
NIDN : 0324018902

Mengetahui,

  
**Dr. Bambang Jokonowo, S.Si., M.T.I.**  
Dekan Fakultas Ilmu Komputer

  
**Dr. Ruci Meivanti, M.Kom**  
Ka.Prodi Sistem Informasi

## KATA PENGANTAR

Alhamdulillah, puji dan syukur penulis panjatkan kehadirat Allah Ta'ala. atas limpahan rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan tugas akhir yang berjudul, “Analisis Data Sentimen Terkait Opini Hasil Pemilu 2024 di Media Sosial dengan Komparasi Model Klasifikasi Text Mining” dengan baik.

Penulis menyadari banyak pihak yang membantu dan berkontribusi dalam terselesaikannya Tugas Akhir ini. Segala bentuk bantuan, baik berupa dukungan moril dan materil sangat membantu penulis dalam mengumpulkan semangat dan keinginan untuk menyelesaikan studi. Dengan demikian, penulis mengucapkan terima kasih kepada :

1. Allah SWT karna telah memberikan rahmat, dan karunia-Nya kepada penulis sehingga penulis dapat bertahan hingga mendapatkan gelar sarjana.
2. Ibu Riri Fajriah, S.Kom., MM selaku dosen pembimbing yang telah memberikan bimbingan, bantuan, serta saran sehingga penulis dapat menyelesaikan laporan ini.
3. Ibu Dr. Ruci Meiyanti, S.Kom, M.Kom selaku Kepala Program Studi Sistem Informasi dan Dosen Pengampu
4. Ibu Nia Rahma Kurnianda, S.kom, M.kom Selaku Dosen Pembimbing Akademik
5. Kedua orangtua Bapak Suaib dan Ibu Syahriawan, Terima kasih karna tiada lelah untuk senantiasa mengirimkan doa, memberikan semangat serta motivasi sehingga penulis mampu menyelesaikan studinya sampai sarjana.
6. Terakhir untuk diri saya sendiri, Indah Fatimah Afriani. Terima kasih telah bertanggung jawab menyelesaikan apa yang telah di mulai. Terima kasih telah bertahan dan senantiasa menikmati proses demi proses yang harus di jalanani.

Akhir kata, penulis berharap segala bentuk saran serta masukan bahkan kritik yang membangun dari berbagai pihak. Semoga laporan tugas akhir ini dapat bermanfaat bagi semua pihak dan khususnya dalam bidang teknologi informasi.

Jakarta, 31 Desember 2024

Indah Fatimah Afriani



## HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademik Universitas Mercu Buana, saya yang bertanda tangan di bawah ini:

Nama : Indah Fatimah Afriani  
NIM : 41820120034  
Pembimbing TA : Riri Fajriah, S.Kom, MM  
Judul : Analisis Data Sentimen Terkait Opini Hasil Pemilu 2024 di Media Sosial dengan Komparasi Model Klasifikasi Text Mining

Demi pengembangan ilmu pengetahuan, dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Non-Eksklusif (*Non-exclusive Royalty-Free Right*)** atas karya ilmiah saya yang berjudul di atas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti Non-Eksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan Laporan Magang/Skripsi/Tesis/Disertasi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 18 Januari 2025

Yang menyatakan,

  
44AMX128/97194  
( Indah Fatimah Afriani )

## ABSTRAK

Nama : Indah Fatimah Afriani  
NIM : 41820120034  
Pembimbing TA : Riri Fajriah, S.Kom, MM  
Judul : Analisis Data Sentimen Terkait Opini Hasil Pemilu 2024 di Media Sosial dengan Komparasi Model Klasifikasi Text Mining

Dalam era digital, media sosial telah menjadi platform populer bagi masyarakat untuk menyampaikan pendapat mereka tentang pemilu. Penelitian ini bertujuan untuk menganalisis data sentimen terkait opini hasil pemilu 2024 di media sosial dengan menggunakan teknik crawling data. Data opini masyarakat diambil dari platform media sosial seperti Twitter, Facebook, dan Instagram. Penggunaan media sosial di Indonesia meningkat pesat, dengan 167 juta pengguna pada tahun 2023, di mana 153 juta adalah pengguna di atas usia 18 tahun. Instagram menjadi platform yang sangat populer, terutama di kalangan anak muda, yang sering membahas isu politik dan pemilu di sana. Data dikumpulkan menggunakan teknik crawling yang memungkinkan pengambilan data secara otomatis. Pada tahap pra-pemrosesan, dilakukan pembersihan data, penghapusan tautan, penghilangan karakter khusus, dan tokenisasi. Selanjutnya, model klasifikasi text mining seperti Naive Bayes dan Random Forest diterapkan untuk mengklasifikasikan opini masyarakat menjadi tiga kategori sentimen: positif, negatif, dan netral. Hasil analisis menunjukkan bahwa teknik crawling data memungkinkan pengumpulan data opini masyarakat secara efisien dan mendalam dari media sosial.

Kata kunci: Analisis Sentimen, Opini Masyarakat, Media Sosial, Crawling Data, Text Mining.



## ABSTRACT

*Name* : Indah Fatimah Afriani  
*NIM* : 41820120034  
*TA Counselor* : Riri Fajriah, S.Kom, MM  
*Title* : *Analysis of Sentiment Data Related to Opinions on the 2024 Election Results on Social Media with Comparison of Text Mining Classification Models*

*In the digital age, social media has become a popular platform for people to share their opinions about elections. This study aims to analyze sentiment data related to opinions of the 2024 election results on social media using data crawling techniques. Public opinion data is obtained from social media platforms, which include Twitter, Facebook, and Instagram. The use of social media in Indonesia is experiencing very rapid development, launching from data reportal in 2023. There are a total of 167 million social media users, 153 million are users over the age of 18 years, which is 79.5 percent of the total population in Indonesia. Instagram is one of the popular social media with interesting pictures compared to reading. Therefore, Instagram is more in demand, especially by young people, one of which is used to discuss political issues and elections. Data collection is done by utilizing data crawling techniques that allow automatic data retrieval from these platforms. At the data pre-processing stage, steps such as data cleansing, link removal, special character removal, and tokenization are performed to prepare the data before the analysis process. Furthermore, several classification models in text mining, including Naive Bayes and Random Forest methods are applied to classify public opinion into three sentiment categories: positive, negative, and neutral. The results of the analysis show that the use of data crawling techniques allows efficient and in-depth collection of public opinion data from social media.*

*Keywords: Sentiment Analysis, Community Opinion, Social Media, Data Crawling, Text Mining*

## DAFTAR ISI

<b>HALAMAN PERNYATAAN KARYA SENDIRI</b> .....	<b>iii</b>
<b>LEMBAR PENGESAHAN</b> .....	<b>iv</b>
<b>KATA PENGANTAR</b> .....	<b>v</b>
<b>HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS</b> .....	<b>vii</b>
<b>ABSTRAK</b> .....	<b>viii</b>
<b>ABSTRACT</b> .....	<b>ix</b>
<b>DAFTAR TABEL</b> .....	<b>xiv</b>
<b>DAFTAR LAMPIRAN</b> .....	<b>xv</b>
<b>BAB I</b> .....	<b>1</b>
<b>PENDAHULUAN</b> .....	<b>1</b>
1.1    Latar Belakang .....	1
1.2    Permasalahan.....	4
1.2.1    Identifikasi Masalah .....	4
1.2.2    Rumusan Masalah .....	5
1.3    Tujuan Penelitian.....	5
1.4    Batasan Masalah.....	5
1.5    Manfaat Penelitian .....	6
1.6    Sistematika Penulisan .....	6
<b>BAB II</b> .....	<b>8</b>
<b>TINJAUAN PUSTAKA</b> .....	<b>8</b>
2.1    Teori/Konsep Terkait.....	8
2.1.1    Analisis Sentimen .....	8

2.1.2	Data Mining .....	9
2.1.3	Text Mining.....	10
2.1.4	Algoritma .....	10
2.1.5	Text Pre-Processing.....	11
2.1.6	Stemming .....	12
2.2	Penelitian Terdahulu.....	13
2.3	Analisa Literatur Review .....	28
2.3.1	Compare: .....	28
2.3.2	Contrast: .....	28
<b>BAB III</b>	.....	<b>30</b>
<b>METODE PENELITIAN</b>	.....	<b>30</b>
3.1	Metode CRISP-DM.....	30
3.2	Deskripsi Sumber Data .....	37
3.3	Teknik Pengumpulan Data .....	37
3.4	Diagram Alir Penelitian.....	37
3.5	Jadwal Penelitian.....	41
<b>BAB IV</b>	.....	<b>42</b>
<b>HASIL DAN PEMBAHASAN</b>	.....	<b>42</b>
4.1	Dataset.....	42
4.2	Pembahasan.....	42
4.2.1	Pengumpulan Data .....	42
4.2.2	<i>Cleaning</i> .....	43
4.2.3	Labeling Dataset.....	45
4.2.4	Preprocessing.....	46
4.3	Evaluasi Model.....	49

4.3.1	Hasil Evaluasi Raw Data.....	51
4.3.2	Hasil Evaluasi Oversampling (SMOTE).....	52
4.3.3	Hasil Evaluasi Random Undersampling .....	53
4.4	Analisa dan Pembahasan.....	54
4.4.1	Pembobotan.....	55
4.4.2	Klasifikasi Random Forest.....	59
4.5	Perbandingan Hasil Metode .....	60
4.5.1	Model yang Paling Baik:.....	60
4.5.2	Pengaruh Teknik Balancing .....	60
4.5.3	Kinerja Berdasarkan Model: .....	61
4.5.4	Visualisasi .....	61
<b>BAB V</b>	.....	<b>64</b>
<b>KESIMPULAN DAN SARAN</b>	.....	<b>64</b>
5.1	Kesimpulan .....	64
5.2	Saran.....	64
<b>DAFTAR PUSTAKA</b>	.....	<b>66</b>
<b>LAMPIRAN</b>	.....	<b>69</b>

## DAFTAR GAMBAR

Gambar 1. 1 Grafik Penggunaan Instagram tahun 2024 .....	2
Gambar 1. 2 Grafik Penggunaan Media Sosial .....	3
<i>Gambar 3. 1 Flow Cleansing Dataset.....</i>	<i>31</i>
Gambar 3. 2 Flow Case Folding.....	32
Gambar 3. 3 Flow Stopwords.....	33
Gambar 3. 4 Flow Labeling Dataset.....	34
Gambar 3. 5 Flow Proses Feature Extraction.....	35
Gambar 3. 6 Diagram Alir Business CRISPP-DM.....	37
Gambar 4. 1 Coding Python Pada Crawling Dataset .....	42
Gambar 4. 2 Hasil Crawling Dataset.....	43
Gambar 4. 3 Code Pada Google Colab Untuk Cleaning Dataset .....	44
Gambar 4. 4 Hasil Cleaning Dataset .....	45
Gambar 4. 5 Pada Labeling Dataset .....	46
Gambar 4. 6 Hasil Labeling Dataset .....	46
Gambar 4. 7 Hasil Labeling Dataset .....	47
Gambar 4. 8 coding pyhton tokenizing dataset .....	47
Gambar 4. 9 Hasil Tokenizing Dataset.....	48
Gambar 4. 10 Confusion Matrix Pada Raw Data.....	50
Gambar 4. 11 Confusion Matrix Pada Raw Data Untuk Random Forest.....	51
Gambar 4. 12 Hasil Evaluasi Over Sampling Pada Naïve Bayes dan Random Forest .....	52
Gambar 4. 13 Hasil Random Undersampling Pada Naïve Bayes .....	53
Gambar 4. 14 Hasil Random Undersampling Pada Random Forest .....	54
Gambar 4. 15 Hasil Confusion Matrix dengan Random Forest Menggunakan SMOTE.....	62
Gambar 4. 16 Confusion Matrix Heatmap pada Random Forest Random Undersampling .....	63

## DAFTAR TABEL

Tabel 2. 2 Literature Review .....	13
Tabel 3. 2 Jadwal Penelitian .....	41
Tabel 4. 1 Labeling Dataset .....	46
Tabel 4. 2 Hasil Tokenizing Dataset .....	48
Tabel 4. 3 Confusion Matrix .....	54



## DAFTAR LAMPIRAN

Lampiran 1 Kartu Asistensi Bimbingan.....	69
Lampiran 2 Surat Pernyataan HKI.....	70
Lampiran 3 Curriculum Vitae.....	71
Lampiran 4 Surat Keterangan BNSP .....	72
Lampiran 5 Hasil Cek Uji Plagiasi.....	73

