

**ANALISIS SENTIMEN MASYARAKAT TERHADAP
ELEKTABILITAS ANIES BASWEDAN SEBAGAI
CALON PRESIDEN PADA PEMILU 2024
DI *TWITTER***

SKRIPSI

**Oleh:
Lutfi Fajri Akbar
201910225014**



**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS BHAYANGKARA JAKARTA RAYA
2023**

LEMBAR PERSETUJUAN PEMBIMBING

LEMBAR PERSETUJUAN PEMBIMBING (DIGUNAKAN UNTUK TUGAS AKHIR)

Judul Tugas Akhir : Analisis Sentimen Masyarakat terhadap Elektabilitas
Anies Baswedan sebagai Calon Presiden pada Pemilu

2024 di *Twitter*

Nama Mahasiswa : Lutfi Fajri Akbar

Nomor Pokok Mahasiswa : 201910225014

Program Studi/Fakultas : Informatika/Ilmu Komputer

Tanggal Lulus Ujian Skripsi : 31 Juli 2023

Jakarta, 10 Agustus 2023

MENYETUJUI

Pembimbing I

Pembimbing II


Prima Dina Atika, S.Kom., M.Kom.
NIDN: 0311037107


Mayadi, S.Kom., M.Kom.
NIDN: 0408087802

LEMBAR PENGESAHAN

LEMBAR PENGESAHAN

Judul Skripsi : Analisis Sentimen Masyarakat Terhadap Elektabilitas
Anies Baswedan sebagai Calon Presiden pada
Pemilu 2024 di *Twitter*

Nama Mahasiswa : Lutfi Fajri Akbar
Nomor Induk Mahasiswa : 201910225014
Program Studi/Fakultas : Informatika/Illmu Komputer
Tanggal Lulus Ujian Skripsi : 31 Juli 2023
Jakarta, 09 Agustus 2023

MENGESAHKAN

Ketua Tim Penguji : Rakhmat Purnomo, S.Pd., S.Kom., M.Kom.
NIDN: 0322108201

Penguji II : Khairunnisa Fadhillah Ramdhanita, M.Si.
NIDN: 0328039201

Penguji III : Prima Dina Atika, S.Kom., M.Kom.
NIDN: 0311037107

MENGETAHUI,

Ketua
Program Studi Informatika

Dekan
Fakultas Ilmu Komputer

Ahmad Fathurrozi, S.E., M.M.S.I.
NIP: 2012486

Dr. Drs. Tyastuti Sri Lestari, M.M.
NIP: 1408206

LEMBAR PERNYATAAN BUKAN PLAGIASI



UNIVERSITAS BHAYANGKARA JAKARTA RAYA
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA

LEMBAR PERNYATAAN BUKAN PLAGIASI

Yang bertanda tangan dibawah ini :

Nama : Lutfi Fajri Akbar
NPM : 201910225014
Program Studi : Informatika
Fakultas : Ilmu Komputer
Judul Tugas Akhir : Analisis Sentimen Masyarakat terhadap Elektabilitas Anies
Baswedan sebagai Calon Presiden pada Pemilu 2024 di
Twitter

Dengan ini menyatakan bahwa hasil penulisan skripsi yang telah saya buat ini merupakan **hasil karya saya sendiri dan benar keasliannya**. Apabila dikemudian hari penulisan skripsi ini merupakan plagiat atau penjiplakan terhadap karya orang lain, maka saya bersedia mempertanggungjawabkan sekaligus bersedia menerima sanksi berdasarkan tata tertib di Universitas Bhayangkara Jakarta Raya.

Demikian pernyataan ini saya buat dalam keadaan sadar dan tidak dipaksakan dari pihak manapun.

Jakarta, 10 Agustus 2023

Penulis


AA3EAKX562966348
Lutfi Fajri Akbar

ABSTRAK

Lutfi Fajri Akbar. 201910225014. Analisis Sentimen Masyarakat terhadap Elektabilitas Anies Baswedan sebagai Calon Presiden pada Pemilu 2024 di *Twitter*. 2023.

Indonesia adalah negara demokrasi yang mengadakan pemilu setiap lima tahun sekali. Anies Baswedan adalah salah satu tokoh politik dengan tingkat elektabilitas tertinggi pada berbagai hasil survei calon presiden pemilu 2024. Adanya media berita *online* yang mempublikasikan hasil survei mengenai elektabilitas Anies Baswedan, membuat masyarakat menyampaikan beragam pendapatnya dalam bentuk *tweets* di *Twitter*. Diperlukan metode *text mining* untuk menganalisis sentimen yang terkandung pada *tweets* tersebut, untuk diketahui polaritas sentimen yang ada pada masyarakat. Metode *text mining* yang digunakan penelitian ini adalah metode klasifikasi *Naïve Bayes* dan *Support Vector Machine (SVM)*. Tujuan penelitian ini yaitu melakukan pengujian metode klasifikasi *Naïve Bayes* dan *SVM* dalam mengklasifikasikan sentimen menggunakan data uji dengan berbagai rasio pembagian data, melakukan perbandingan hasil nilai akurasi yang didapatkan kedua metode *text mining* dengan metrik *accuracy*, dan mengetahui polaritas sentimen masyarakat berdasarkan hasil analisis sentimen metode *text mining* terbaik. Hasil penelitian ini adalah metode klasifikasi *SVM* pada persentase rasio pembagian data 90:10 merupakan metode *text mining* dengan performa terbaik dalam melakukan *text mining* dengan nilai akurasi sebesar 86% dan mayoritas sentimen masyarakat cenderung berpolaritas sentimenkan negatif terhadap topik elektabilitas Anies Baswedan sebagai calon presiden pada pemilu 2024, dengan persentase sentimen negatif mencapai 56.9%.

Kata Kunci: Analisis Sentimen, *Text Mining*, Metode Klasifikasi, *Naïve Bayes*, *SVM*

ABSTRACT

Lutfi Fajri Akbar. 201910225014. *Analysis of Public Sentiment towards Anies Baswedan's Electability as a Presidential Candidate in the 2024 Election on Twitter. 2023.*

Indonesia is a democratic country that holds elections every five years. Anies Baswedan is one of the political figures with the highest level of electability in various surveys of presidential candidates for the 2024 elections. The existence of online news media that publishes survey results regarding Anies Baswedan's electability, makes people express various opinions in the form of tweets on Twitter. A text mining method is needed to analyze the sentiment contained in the tweets, to find out the polarity of sentiment that exists in the community. The text mining methods used in this research are Naïve Bayes and Support Vector Machine (SVM) classification methods. The purpose of this research is to test the Naïve Bayes and SVM classification methods in classifying sentiment using test data with various data division ratios, comparing the results of the accuracy values obtained by the two text mining methods with the accuracy metric, and knowing the polarity of community sentiment based on the results of the sentiment analysis of the best text mining method. The result of this research is that the SVM classification method on the percentage of data division ratio 90:10 is the text mining method with the best performance in conducting text mining with an accuracy value of 86% and the majority of public sentiment tends to have negative sentiment polarity on the topic of Anies Baswedan's electability as a presidential candidate in the 2024 election, with a percentage of negative sentiment reaching 56.9%.

Keywords: Sentiment Analysis, Text Mining, Classification Method, Naïve Bayes, SVM

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIK

Sebagai sivitas akademik Universitas Bhayangkara Jakarta Raya, saya yang bertanda tangan di bawah ini :

Nama : Lutfi Fajri Akbar
NPM : 201910225014
Program Studi : Informatika
Fakultas : Ilmu Komputer
Jenis Karya : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Bhayangkara Jakarta Raya **Hak Bebas Royalti Non-Eksklusif (Non-Exclusive Royalty-Free Right)**, atas karya ilmiah saya yang berjudul :

Analisis Sentimen Masyarakat terhadap Elektabilitas Anies Baswedan sebagai Calon Presiden pada Pemilu 2024 di *Twitter*

beserta perangkat yang ada (bila diperlukan). Dengan hak bebas royalti non-eksklusif ini, Universitas Bhayangkara Jakarta Raya berhak menyimpan, mengalihmediakan, mengelolanya dalam bentuk pangkalan data (*database*), mendistribusikannya dan mempublikasikannya di Internet atau media lain untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis dan sebagai pemilik hak cipta.

Segala bentuk tuntutan hukum yang timbul atas pelanggaran hak cipta dalam karya ilmiah ini menjadi tanggung jawab saya pribadi

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta
Pada tanggal : 10 Agustus 2023
Yang Menyatakan


Lutfi Fajri Akbar

KATA PENGANTAR

Dengan memanjatkan puja dan puji syukur atas kehadiran Allah SWT yang telah melimpahkan rahmat, taufik, dan hidayah-Nya sehingga skripsi dengan judul “Analisis Sentimen Masyarakat terhadap Elektabilitas Anies Baswedan sebagai Calon Presiden pada Pemilu 2024 di *Twitter*” ini dapat terselesaikan. Skripsi ini ditulis untuk memenuhi persyaratan untuk menyelesaikan pendidikan Strata Satu (S1) pada Program Studi Informatika Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya.

Disadari bahwa skripsi ini tidak mungkin dapat terselesaikan tanpa adanya bantuan, dukungan, bimbingan, dan nasehat dari berbagai pihak selama penyusunan skripsi ini. Oleh karena itu, diucapkannya kata terima kasih setulus-tulusnya kepada kedua orang tua yang telah mendoakan serta memberikan dukungan dalam penulisan skripsi ini dan kata terima kasih diucapkan juga kepada:

1. Bapak Irjen Polisi (Purn) Dr. Drs. H. Bambang Karsono, S.H., M.M. selaku Rektor Universitas Bhayangkara Jakarta Raya;
2. Ibu Dr. Dra. Tyastuti Sri Lestari, M.M. selaku Dekan Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya;
3. Bapak Ahmad Fathurrozi, S.E., M.M.S.I. selaku Ketua Program Studi Informatika;
4. Ibu Prima Dina Atika, S.Kom., M.Kom. selaku dosen pembimbing I penyusunan skripsi;
5. Bapak Mayadi, S.Kom., M.Kom. selaku dosen pembimbing II penyusunan skripsi;

6. Bapak Rakhmat Purnomo, S.Pd., S.Kom., M.Kom. selaku dosen penguji I sidang skripsi;
7. Ibu Khairunnisa Fadhilla Ramdhanian, S.Si., M.Si. selaku dosen penguji II sidang skripsi;
8. Bapak Muhammad Yasir, S.Si., M.Kom. selaku mentor di bidang *data mining*;
9. Ibu Siti Setiawati, S.Pd., M.Pd. selaku validator ahli bahasa proses *labelling* data;
10. Teman-teman yang sudah membantu dan memberi dukungan dalam penulisan skripsi ini.

Disadari bahwa skripsi ini jauh dari kesempurnaan, untuk itu mohon maaf apabila terdapat kekurangan dan kesalahan baik dari pembahasan maupun penulisan. Oleh karena itu, diharapkan masukan berupa kritik dan saran yang membangun.

Bekasi, 10 Agustus 2023



Lutfi Fajri Akbar

201910225014

DAFTAR ISI

LEMBAR PERSETUJUAN PEMBIMBING	i
LEMBAR PENGESAHAN	ii
LEMBAR PERNYATAAN BUKAN PLAGIASI	iii
ABSTRAK	iv
ABSTRACT	v
LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI	vi
KATA PENGANTAR.....	vii
DAFTAR ISI.....	ix
DAFTAR TABEL	xii
DAFTAR GAMBAR.....	xiii
DAFTAR LAMPIRAN	xv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Identifikasi Masalah.....	5
1.3 Rumusan Masalah.....	5
1.4 Batasan Masalah	6
1.5 Tujuan Penelitian	6
1.6 Manfaat Penelitian	7
1.7 Sistematika Penulisan	7
BAB II TINJAUAN PUSTAKA.....	9
2.1 Penelitian Relevan	9
2.1.1 Tabel Related Research / State of The Art	13
2.2 Landasan Teori.....	16
2.2.1 Pemilu	16
2.2.2 Elektabilitas	16
2.2.3 Opini	17
2.2.4 Twitter	18
2.2.5 Analisis Sentimen	19
2.2.6 Data Mining.....	20
2.2.7 Text Mining	21
2.2.8 Text Preprocessing	23
2.2.9 Metode Naïve Bayes	27

2.2.10	Metode <i>Support Vector Machine</i>	30
2.2.11	<i>Confusion Matrix</i>	33
2.2.12	<i>CRISP-DM</i>	35
2.2.12	<i>Python</i>	38
BAB III METODOLOGI PENELITIAN		39
3.1	Kerangka Penelitian	39
3.1.1	<i>Business Understanding</i>	40
3.1.2	<i>Data Understanding</i>	42
3.1.3	<i>Data Preparation</i>	42
3.1.4	<i>Modelling</i>	43
3.1.5	<i>Evaluation</i>	43
3.1.6	<i>Result</i>	44
3.1.7	Analisis Kebutuhan <i>Hardware</i> dan <i>Software</i>	44
BAB IV HASIL DAN PEMBAHASAN		46
4.1	Hasil <i>Data Understanding</i>	46
4.2	Hasil <i>Data Preparation</i>	50
4.2.1	Hasil Penghapusan <i>Missing Value</i> dan <i>Duplicate Data</i>	50
4.2.2	Hasil <i>Labelling</i>	52
4.2.3	Hasil <i>Label Encoder</i>	54
4.2.4	Hasil <i>Case Folding</i>	55
4.2.5	Hasil <i>Cleansing</i>	57
4.2.6	Hasil <i>Tokenization</i>	58
4.2.7	Hasil <i>Normalization</i>	59
4.2.8	Hasil <i>Remove Stopwords</i>	65
4.2.9	Hasil <i>Stemming</i>	67
4.2.10	Hasil <i>Data Splitting</i>	68
4.2.11	Hasil Pembobotan Kata Menggunakan Metode <i>TF-IDF</i>	75
4.3	Hasil <i>Modelling</i>	79
4.3.1	Hasil <i>Modelling</i> Metode <i>Naïve Bayes</i>	79
4.3.2	Hasil <i>Modelling</i> Metode <i>SVM</i>	84
4.4	Hasil <i>Evaluation</i>	88
4.4.6	Perbandingan Akurasi Model Metode <i>Text Mining</i>	104
4.4.7	Uji Prediksi Model Metode <i>Text Mining</i> Terbaik	106
4.5	<i>Result</i> Penelitian	107

BAB V PENUTUP.....	110
5.1 Kesimpulan	110
5.2 Saran	111
DAFTAR PUSTAKA.....	112
LAMPIRAN.....	116



DAFTAR TABEL

Tabel 2.1 <i>Related Research / State of The Art</i>	13
Tabel 2.2 Tabel <i>Confusion Matrix</i>	33
Tabel 3.1 Spesifikasi <i>Hardware</i> Milik <i>Google Colaboratory</i>	45
Tabel 4.1 Hasil <i>Labelling Data</i>	53
Tabel 4.2 Hasil <i>Label Encoder</i>	55
Tabel 4.3 Hasil <i>Case Folding</i>	56
Tabel 4.4 Hasil <i>Cleansing</i>	57
Tabel 4.5 Hasil <i>Tokenization</i>	59
Tabel 4.6 Daftar Kata Kamus <i>Slang Words</i>	59
Tabel 4.7 Hasil <i>Normalization</i>	64
Tabel 4.8 Hasil <i>Remove Stopwords</i>	66
Tabel 4.9 Hasil <i>Stemming</i>	68
Tabel 4.10 Persentase Rasio Pembagian Data	69
Tabel 4.11 Lima Dokumen Penghitungan Pembobotan Kata Metode <i>TF-IDF</i>	76
Tabel 4.12 Hasil Penghitungan Manual Rumus Metode <i>TF-IDF</i>	77
Tabel 4.13 Hasil Pembobotan Kata Metode <i>TF-IDF</i> pada Kelima Dokumen.....	78
Tabel 4.14 Lima Dokumen Penghitungan Model <i>Multinomial Naïve Bayes</i>	81
Tabel 4.15 Lima Dokumen Penghitungan Model <i>SVM</i>	86
Tabel 4.16 Tabel <i>Confusion Matrix</i> Model <i>Naïve Bayes</i> Persentase Rasio 90:10.....	91
Tabel 4.17 Tabel <i>Confusion Matrix</i> Model <i>SVM</i> Persentase Rasio 90:10.....	92
Tabel 4.18 Tabel <i>Confusion Matrix</i> Model <i>Naïve Bayes</i> Persentase Rasio 80:20.....	93
Tabel 4.19 Tabel <i>Confusion Matrix</i> Model <i>SVM</i> Persentase Rasio 80:20.....	95
Tabel 4.20 Tabel <i>Confusion Matrix</i> Model <i>Naïve Bayes</i> Persentase Rasio 70:30.....	96
Tabel 4.21 Tabel <i>Confusion Matrix</i> Model <i>SVM</i> Persentase Rasio 70:30.....	98
Tabel 4.22 Tabel <i>Confusion Matrix</i> Model <i>Naïve Bayes</i> Persentase Rasio 60:40.....	99
Tabel 4.23 Tabel <i>Confusion Matrix</i> Model <i>SVM</i> Persentase Rasio 60:40.....	101
Tabel 4.24 Tabel <i>Confusion Matrix</i> Model <i>Naïve Bayes</i> Persentase Rasio 50:50...	102
Tabel 4.25 Tabel <i>Confusion Matrix</i> Model <i>SVM</i> Persentase Rasio 50:50.....	104
Tabel 4.26 Perbandingan Nilai Akurasi Kedua Model Metode	104
Tabel 4.27 Daftar Data Uji Baru	106

DAFTAR GAMBAR

Gambar 2.1 Proses <i>Text Mining</i>	22
Gambar 2.2 Ilustrasi Metode SVM.....	31
Gambar 2.3 <i>Small Margin</i> dan <i>Large Margin</i>	32
Gambar 2.4 Siklus Hidup Proyek <i>Data Mining CRISP-DM</i>	35
Gambar 2.5 Hasil Survei KDnuggets.....	37
Gambar 3.1 Kerangka Penelitian	40
Gambar 4.1 Fase Pertama <i>Scraping Data Tweets</i>	46
Gambar 4.2 Fase Kedua <i>Scraping Data Tweets</i>	47
Gambar 4.3 Fase Ketiga <i>Scraping Data Tweets</i>	47
Gambar 4.4 <i>Missing Value</i> yang Ditemui pada Data.....	48
Gambar 4.5 <i>Duplicate Data</i> yang Ditemui pada Data Teks <i>Tweets</i>	49
Gambar 4.6 Kata Tidak Baku dan Kata <i>Slang</i> yang Ditemui pada Data	49
Gambar 4.7 Hal-Hal yang Dapat Menjadikan <i>Noise</i> pada Data	50
Gambar 4.8 Hasil Penghapusan <i>Missing Value</i>	51
Gambar 4.9 Hasil Penghapusan <i>Duplicate Data</i>	51
Gambar 4.10 Hasil <i>Labelling Data</i>	53
Gambar 4.11 Perbandingan Jumlah Data Kelas <i>Positive</i> dan <i>Negative</i>	54
Gambar 4.12 Kode <i>Python Label Encoder</i>	54
Gambar 4.13 Kode <i>Python Case Folding</i>	56
Gambar 4.14 Kode <i>Python Cleansing</i>	57
Gambar 4.15 Kode <i>Python Tokenization</i>	58
Gambar 4.16 Kode <i>Python Normalization</i>	64
Gambar 4.17 Kode <i>Python Remove Stopwords</i>	66
Gambar 4.18 Kode <i>Python Stemming</i>	67
Gambar 4.19 Kode <i>Python</i> Persentase Rasio Pembagian 90:10	69
Gambar 4.20 Perbandingan Jumlah Data Latih dan Data Uji Rasio 90:10.....	70
Gambar 4.21 Kode <i>Python</i> Persentase Rasio Pembagian 80:20	70
Gambar 4.22 Perbandingan Jumlah Data Latih dan Data Uji Rasio 80:20.....	71
Gambar 4.23 Kode <i>Python</i> Persentase Rasio Pembagian 70:30	71
Gambar 4.24 Perbandingan Jumlah Data Latih dan Data Uji Rasio 70:30.....	72
Gambar 4.25 Kode <i>Python</i> Persentase Rasio Pembagian 60:40	72
Gambar 4.26 Perbandingan Jumlah Data Latih dan Data Uji Rasio 60:40.....	73

Gambar 4.27 Kode <i>Python</i> Persentase Rasio Pembagian 50:50	74
Gambar 4.28 Perbandingan Jumlah Data Latih dan Data Uji Rasio 50:50.....	74
Gambar 4.29 Kode <i>Python TF-IDF Vectorizer</i>	75
Gambar 4.30 <i>Define Model</i> Metode <i>Naïve Bayes</i>	79
Gambar 4.31 <i>Training Model</i> Metode <i>Naïve Bayes</i>	79
Gambar 4.32 <i>Predict Model</i> Metode <i>Naïve Bayes</i>	80
Gambar 4.33 <i>Define Model</i> Metode <i>SVM</i>	84
Gambar 4.34 <i>Training Model</i> Metode <i>SVM</i>	85
Gambar 4.35 <i>Predict Model</i> Metode <i>SVM</i>	85
Gambar 4.36 Kode <i>Python Plot Confusion Matrix</i> Model <i>Naïve Bayes</i>	89
Gambar 4.37 Kode <i>Python Plot Confusion Matrix</i> Model <i>SVM</i>	89
Gambar 4.38 <i>Plot Confusion Matrix</i> Model <i>Naïve Bayes</i> Persentase Rasio 90:10 ...	90
Gambar 4.39 <i>Plot Confusion Matrix</i> Model <i>SVM</i> Persentase Rasio 90:10.....	91
Gambar 4.40 <i>Plot Confusion Matrix</i> Model <i>Naïve Bayes</i> Persentase Rasio 80:20 ...	93
Gambar 4.41 <i>Plot Confusion Matrix</i> Model <i>SVM</i> Persentase Rasio 80:20.....	94
Gambar 4.42 <i>Plot Confusion Matrix</i> Model <i>Naïve Bayes</i> Persentase Rasio 70:30 ...	96
Gambar 4.43 <i>Plot Confusion Matrix</i> Model <i>SVM</i> Persentase Rasio 70:30.....	97
Gambar 4.44 <i>Plot Confusion Matrix</i> Model <i>Naïve Bayes</i> Persentase Rasio 60:40 ...	99
Gambar 4.45 <i>Plot Confusion Matrix</i> Model <i>SVM</i> Persentase Rasio 60:40.....	100
Gambar 4.46 <i>Plot Confusion Matrix</i> Model <i>Naïve Bayes</i> Persentase Rasio 50:50 .	102
Gambar 4.47 <i>Plot Confusion Matrix</i> Model <i>SVM</i> Persentase Rasio 50:50.....	103
Gambar 4.48 Kode <i>Python</i> Pengujian Prediksi Sentimen dengan Data Uji Baru....	106
Gambar 4.49 Hasil Pengujian Prediksi Sentimen dengan Data Uji Baru	107
Gambar 4.50 Kode <i>Python</i> Diagram Lingkaran Polaritas Sentimen	108
Gambar 4.51 Diagram Lingkaran Polaritas Sentimen Masyarakat.....	109

DAFTAR LAMPIRAN

Lampiran 1: Uji Plagiasi	116
Lampiran 2: Biodata Peneliti	117
Lampiran 3: Kartu Konsultasi Bimbingan	118
Lampiran 4: Rekomendasi dari Pembimbing untuk Tugas Akhir	121
Lampiran 5: Lembar Validasi Instrumen Pertama	122
Lampiran 6: Lembar Validasi Instrumen Kedua.....	128
Lampiran 7: Surat Pernyataan Validitas Instrumen	134

