

**ANALISIS SENTIMEN TERHADAP PEMILU 2024
MENGUNAKAN ALGORITMA MULTINOMIAL NAÏVE
BAYES**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Sistem Informasi



disusun oleh

FISKA YOGI HAMDANI

21.12.2159

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2025**

**ANALISIS SENTIMEN TERHADAP PEMILU 2024
MENGUNAKAN ALGORITMA MULTINOMIAL NAÏVE
BAYES**

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Sistem Informasi



disusun oleh

FISKA YOGI HAMDANI

21.12.2159

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

HALAMAN PERSETUJUAN

SKRIPSI

**ANALISIS SENTIMEN TERHADAP PEMILU 2024 MENGGUNAKAN
ALGORITMA MULTINOMIAL NAÏVE BAYES**

yang disusun dan diajukan oleh

Fiska Yogi Hamdani

21.12.2159

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 21 Agustus 2025

Dosen Pembimbing,



Norhikmah, S.Kom., M.Kom
NIK. 190302245

HALAMAN PENGESAHAN

SKRIPSI

**ANALISIS SENTIMEN TERHADAP PEMILU 2024 MENGGUNAKAN
ALGORITMA MULTINOMINAL NAÏVE BAYES**

yang disusun dan diajukan oleh

Fiska Yogi Hamdani

21.12.2159

Telah dipertahankan di depan Dewan Penguji
pada tanggal 21 Agustus 2025

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Bety Wulan Sari, S.Kom., M.Kom.
NIK. 190302254

Alfie Nur Rahmi, S.kom., M.kom.
NIK. 190302240

Norhikmah, S.Kom., M.Kom
NIK. 190302245



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 21 Agustus 2025

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusriul, S.Kom., M.Kom.
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Fiska Yogi Hamdani
NIM : 21.12.2159

Menyatakan bahwa Skripsi dengan judul berikut:

**ANALISIS SENTIMEN TERHADAP PEMILU 2024 MENGGUNAKAN
ALGORITMA MULTINOMINAL NAÏVE BAYES**

Dosen Pembimbing : Norhikmah, S.Kom., M.Kom

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, 21 Agustus 2025

Yang Menyatakan,


Fiska Yogi Hamdani

HALAMAN PERSEMBAHAN

Alhamdulillah rabbil ‘alamin, puji syukur penulis panjatkan ke hadirat Allah SWT atas limpahan rahmat, nikmat, dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi ini dengan baik. Dengan penuh rasa syukur dan cinta, karya ini penulis persembahkan kepada:

1. Kedua orang tua tercinta, Bapak Fais Solichan dan Ibu Sukati, yang senantiasa memberikan cinta, doa, dukungan, dan pengorbanan tiada henti. Terima kasih atas setiap tetesan keringat, motivasi, dan kasih sayang yang menjadi kekuatan penulis dalam melewati segala rintangan hingga akhirnya dapat menuntaskan studi ini.
2. Ibu Norhikmah, S.Kom., M.Kom, selaku dosen pembimbing, yang telah dengan sabar memberikan arahan, nasihat, dan motivasi hingga penulis mampu menyelesaikan skripsi ini.
3. Seluruh dosen Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta, yang telah membagikan ilmu, bimbingan, dan inspirasi selama masa perkuliahan.
4. Untuk sahabat-sahabat yang menjadi bagian tak terpisahkan dari kisah ini, terima kasih atas kebersamaan, dukungan, dan tawa yang senantiasa menghidupkan hari-hari penulis. Semoga persahabatan ini tetap abadi hingga akhir cerita.
5. Seseorang yang istimewa, Kiki Agustina, yang dengan setia memberikan dukungan, menjadi pendengar di setiap keluh kesah, serta menemani penulis melalui setiap tahap dalam penyusunan skripsi ini.
6. Kepada seluruh pihak yang tidak dapat penulis sebutkan satu per satu, terima kasih atas setiap doa, bantuan, dan dukungan yang telah diberikan. Semoga segala kebaikan tersebut mendapatkan balasan berlipat ganda dari Allah SWT.

KATA PENGANTAR

Puji syukur penulis panjatkan kehadiran Tuhan Yang Maha Esa, yang telah melimpahkan rahmat, hidayah, serta karunia-Nya, sehingga penulis dapat menyelesaikan laporan tugas akhir ini dengan lancar. Tugas akhir ini disusun untuk memenuhi salah satu syarat kelulusan Program Sarjana pada Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.

Selain itu penulis dengan segala kerendahan hati ingin menyampaikan rasa terima kasih kepada semua pihak yang telah berjasa memberikan dukungan dan bantuan untuk menyelesaikan laporan tugas akhir ini. Untuk itu penulis mengucapkan terima kasih kepada :

1. Bapak Prof. Dr. M. Suyanto, M.M selaku Rektor Universitas Amikom Yogyakarta
2. Ibu Prof. Dr. Kusrini, M. Kom selaku Dekan Program Fakultas Ilmu Komputer
3. Bapak Anggit Dwi Hartono, M.Kom selaku Ketua Program Studi Sistem Informasi
4. Ibu Norhikmah, S.Kom., M.Kom selaku dosen pembimbing yang memberikan arahan, saran, dan motivasi terhadap penulis
5. Kedua orang tua, keluarga besar serta teman-teman tercinta yang memberikan semangat,dukungan,kasih sayang dan doa kepada penulis.
6. Semua pihak yang tidak dapat disebutkan satu persatu yang telah membantu dalam penyelesaian tugas akhir ini

Yogyakarta, 21 Agustus 2025

Fiska Yogi Hamdani

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iiiv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI.....	vii
DAFTAR TABEL.....	x
DAFTAR GAMBAR	xi
DAFTAR LAMBANG DAN SINGKATAN	xii
DAFTAR ISTILAH	xiii
INTISARI	xiv
ABSTRACT.....	xv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	5
1.3 Batasan Masalah	5
1.4 Tujuan Penelitian	5
1.5 Manfaat Penelitian	6
1.6 Sistematika Penulisan	6
BAB II TINJAUAN PUSTAKA	8
2.1 Studi Literatur	8
2.2 Dasar Teori	19
2.2.1 Pemilihan Umum(Pemilu)	19
2.2.2 Twitter(X)	19
2.2.3 Analisis Sentimen	19

2.2.4 Text Mining	20
2.2.5 Crawling.....	20
2.2.6 Preprocessing	20
2.2.7 .Pelabelan Data	22
2.2.8. TF-IDF	22
2.2.9 SMOTE	23
2.3. Naïve Bayes.....	24
2.4 Mutlinominal Naïve Bayes.....	24
2.5 Confusion Matrix	25
2.6 Visualization.....	26
BAB III METODE PENELITIAN	27
3.1 Objek Penelitian.....	27
3.2 Alur Penelitian	27
3.2.1 Crawling.....	29
3.2.2 Pre-Processing.....	29
3.2.3 Labeling	31
3.2.4. TF-IDF	31
3.2.5. Pembagian Data Latih Dan Uji	31
3.2.6. SMOTE	32
3.2.6. Naïve Bayes	32
3.2.7 Evaluasi.....	32
3.3 Alat dan Bahan.....	33
BAB IV HASIL DAN PEMBAHASAN	34
4.1. <i>Crawling</i> Data	34
4.2. Preprocessing.....	37
4.3 . Pelabelan Data	47
4.4. Pembobotan TF-IDF.....	52
4.5. Pembagian Data Latih dan Data Uji.....	54

4.6. SMOTE	56
4.7 Naïve Bayes.....	58
4.7.1. Multinomial Naïve Bayes	58
4.8 Evaluasi	62
4.9. Implementasi Model.....	62
4.9.1. Halaman Utama	63
4.9.2. Halaman Distribusi Sentimen	64
4.9.3. Halaman pencarian Sentimen	65
4.9.4. Halaman Evaluasi Model	67
BAB V PENUTUP	68
5.1 Kesimpulan	68
5.2 Saran	68
REFERENSI	70

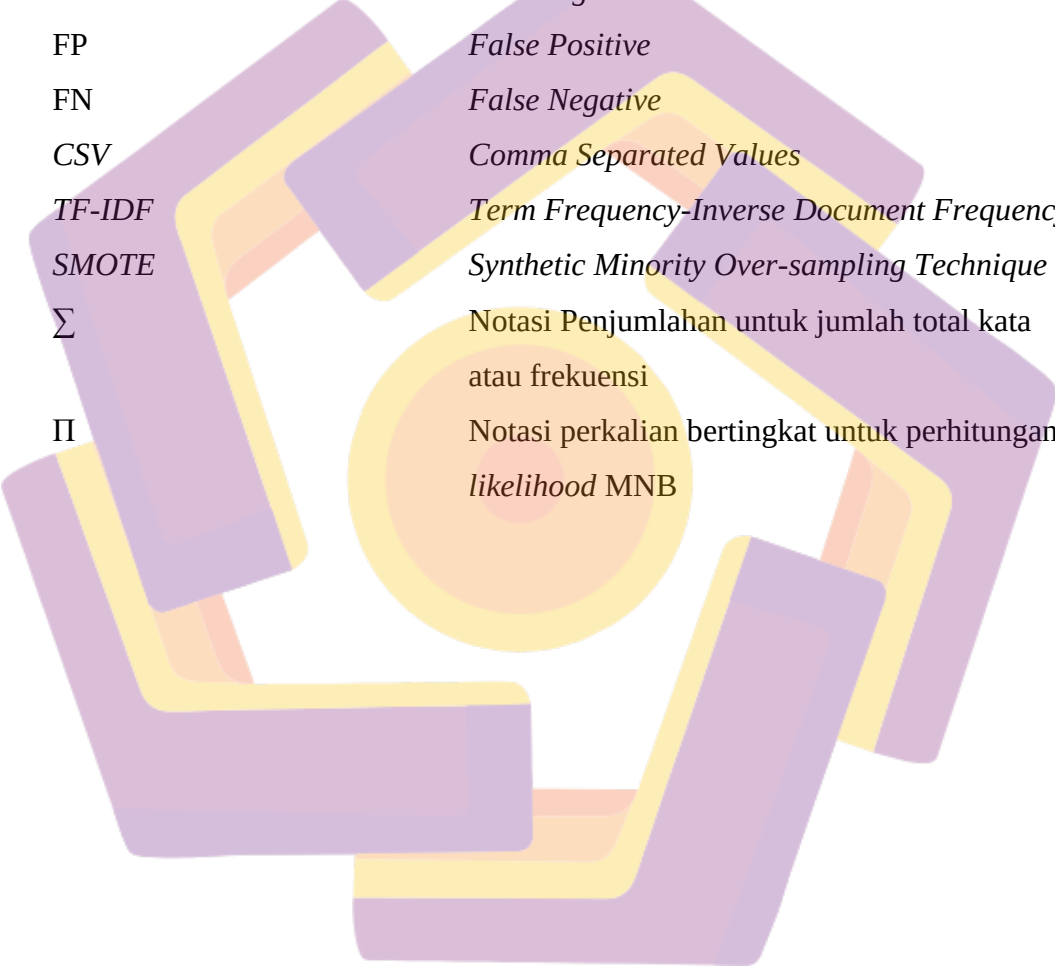
DAFTAR TABEL

Tabel 2.1 Keaslian Penelitian	12
Tabel 2.2 Tabel Confusion Matrix.....	25
Tabel 4.1 Hasil Cleaning.....	38
Tabel 4.2 Hasil Case Folding.....	40
Tabel 4.3 Hasil Case Folding.....	42
Tabel 4.4 Hasil Tokenization.....	43
Tabel 4.5 Stopword Removal.....	45
Tabel 4.6 Hasil dari Stemming	46
Tabel 4.7 Pembobotan TF-IDF	53
Tabel 4.8 Pembagian Data Latih dan Data Uji	56
Tabel 4.9 Hasil SMOTE	57
Tabel 4.10 Confusion Matrix Multinomial NB 90:10	60
Tabel 4.11 Confusion Matrix Multinomial NB 80:20	60
Tabel 4.12 Confusion Matrix Multinomial NB 70:30	61
Tabel 4.13 Evaluasi Kinerja Model Naive Bayes terhadap Klasifikasi Sentimen.....	61

DAFTAR GAMBAR

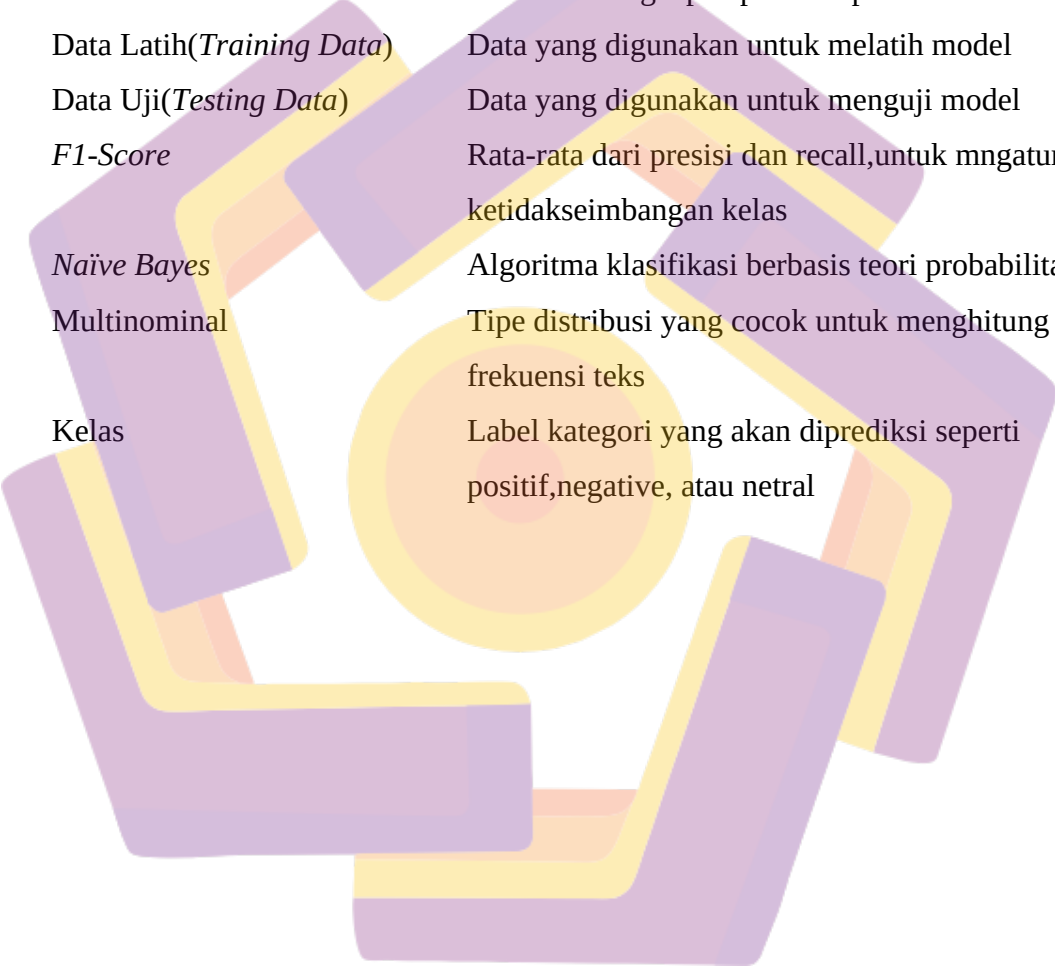
Gambar 3.1 Alur Penelitian	28
Gambar 4.1 Skript Crawling Data	34
Gambar 4.2 Hasil Crawling Data	35
Gambar 4.3 Skript Hapus Duplikat.....	36
Gambar 4.4 Dataset yang dihapus.....	36
Gambar 4.5 Skript Cleaning	37
Gambar 4.6 Skript Case Folding.....	39
Gambar 4.7 Skript Word Normalization.....	41
Gambar 4.8 Skript Tokenization.....	43
Gambar 4.9 Skript Stopword Removal.....	44
Gambar 4.10 Skript Stemming	46
Gambar 4.11 Skript Pelabelan Data.....	47
Gambar 4.12 Hasil dari Pelabelan Data	48
Gambar 4.13 Hasil tampilan Pie Chart	49
Gambar 4.14 Skript dari Word Cloud.....	50
Gambar 4.15 Hasil dari word cloud Sentiment Positif	51
Gambar 4.16 Hasil dari word cloud Sentiment Negatif.....	51
Gambar 4.17 Skript TF-IDF	52
Gambar 4.18 Hasil Visualisasi Bar TF-IDF	54
Gambar 4.19 Skript Split Data.....	55
Gambar 4.20 lanjutan Skript Split Data.....	55
Gambar 4.21 Skript SMOTE	57
Gambar 4.22 Skript Multinomial NB	59
Gambar 4.23 Halaman Utama.....	63
Gambar 4.24 Distribusi Sentimen.....	64
Gambar 4.25 Halaman Pencarian Sentimen	65
Gambar 4.26 Halaman Evaluasi Model	67

DAFTAR LAMBANG DAN SINGKATAN



NB	Naïve Bayes
MNB	Multinomial Naïve Bayes
TF	<i>Term Frequency</i>
TP	<i>True Positive</i>
TN	<i>True Negative</i>
FP	<i>False Positive</i>
FN	<i>False Negative</i>
CSV	<i>Comma Separated Values</i>
<i>TF-IDF</i>	<i>Term Frequency-Inverse Document Frequency</i>
<i>SMOTE</i>	<i>Synthetic Minority Over-sampling Technique</i>
Σ	Notasi Penjumlahan untuk jumlah total kata atau frekuensi
Π	Notasi perkalian bertingkat untuk perhitungan <i>likelihood</i> MNB

DAFTAR ISTILAH



<i>Confusion Matrix</i>	Matrix yang digunakan untuk evaluasi perfroma dari klasifikasi
Akurasi	Ukuran performa model klasifikasi
Presisi	Ukuran ketepatan prediksi positif
<i>Recall</i>	Ukuran kelengkapan prediksi positif
Data Latih(<i>Training Data</i>)	Data yang digunakan untuk melatih model
Data Uji(<i>Testing Data</i>)	Data yang digunakan untuk menguji model
<i>F1-Score</i>	Rata-rata dari presisi dan recall, untuk mengatur ketidakseimbangan kelas
<i>Naïve Bayes</i>	Algoritma klasifikasi berbasis teori probabilitas
Multinomial	Tipe distribusi yang cocok untuk menghitung frekuensi teks
Kelas	Label kategori yang akan diprediksi seperti positif, negative, atau netral

INTISARI

Pemilu 2024 menjadi topik yang ramai diperbincangkan di media sosial, khususnya pada Platform X (Twitter). Opini yang dibagikan mencerminkan pandangan publik terhadap proses demokrasi dan kandidat yang terlibat, namun volume data yang besar membuat analisis manual tidak efisien. Penelitian ini bertujuan untuk menganalisis sentimen masyarakat terhadap Pemilu 2024 secara otomatis sehingga persepsi publik dapat dipahami dengan lebih cepat dan akurat.

Data dikumpulkan melalui teknik *crawling* dari Platform X, kemudian diproses menggunakan tahapan *preprocessing* seperti pembersihan teks, tokenisasi, normalisasi, dan *stemming*. Klasifikasi sentimen dilakukan dengan algoritma Multinomial *Naive Bayes* menggunakan pembobotan *TF-IDF*. Sentimen dibagi menjadi dua kategori, yaitu positif dan negatif. Untuk mengatasi ketidakseimbangan data digunakan metode SMOTE, dan model dievaluasi pada tiga skenario pembagian data latih dan uji (90:10, 80:20, dan 70:30).

Hasil penelitian menunjukkan bahwa skenario 90:10 memberikan performa terbaik dengan akurasi 80,45%, presisi 81,54%, *recall* 79,80%, dan F1-score 80,62%. Temuan ini membuktikan bahwa algoritma Multinomial *Naive Bayes* efektif untuk mengklasifikasikan opini publik, serta berpotensi digunakan sebagai alat bantu berbasis data dalam memahami dinamika persepsi masyarakat menjelang Pemilu 2024.

Kata kunci: Analisis Sentimen, Pemilu 2024, *TF-IDF*, Multinomial *Naive Bayes*, Twitter

ABSTRACT

The 2024 General Election has become a widely discussed topic on social media, particularly on Platform X (Twitter). Public opinions shared online reflect perceptions of the democratic process and the candidates involved, yet the large volume of data makes manual analysis inefficient. This study aims to analyze public sentiment toward the 2024 General Election automatically, enabling a faster and more accurate understanding of public perception.

Data were collected using a crawling technique from Platform X and processed through preprocessing stages such as text cleaning, tokenization, normalization, and stemming. Sentiment classification was performed using the Multinomial Naive Bayes algorithm with TF-IDF weighting. Sentiments were categorized into two classes: positive and negative. To address class imbalance, the SMOTE method was applied, and the model was evaluated under three train-test split scenarios (90:10, 80:20, and 70:30).

The results show that the 90:10 split scenario yielded the best performance, achieving an accuracy of 80.45%, precision of 81.54%, recall of 79.80%, and F1-score of 80.62%. These findings indicate that the Multinomial Naive Bayes algorithm is effective for classifying public opinion and has the potential to serve as a data-driven tool for understanding public sentiment dynamics ahead of the 2024 General Election.

Keyword: Sentiment Analysis, 2024 General Election, TF-IDF, Multinomial Naive Bayes, Twitter