

**ANALISIS SENTIMEN DATA TWITTER TERHADAP
GUGATAN PEMILU 2024 MENGGUNAKAN ALGORITMA
NAIVE BAYES
SKRIPSI**

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



disusun oleh

**MUHAMMAD TAMAMUL ABIL FIDAA
20.11.3463**

Kepada

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

**ANALISIS SENTIMEN DATA TWITTER TERHADAP
GUGATAN PEMILU 2024 MENGGUNAKAN ALGORITMA
NAÏVE BAYES**

SKRIPSI

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



disusun oleh

MUHAMMAD TAMAMUL ABIL FIDAA
20.11.3463

Kepada

FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2025

HALAMAN PERSETUJUAN

SKRIPSI

**ANALISIS SENTIMEN DATA TWITTER TERHADAP
GUGATAN PEMILU 2024 MENGGUNAKAN ALGORITMA
NAÏVE BAYES**

yang disusun dan diajukan oleh

Muhammad Tamamul Abil Fidaa

20.11.3463

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 10 Juni 2025

Dosen Pembimbing,


Mulia Sulistiyono, S.Kom., M.Kom.
NIK. 190302248

HALAMAN PENGESAHAN
SKRIPSI
ANALISIS SENTIMEN DATA TWITTER TERHADAP
GUGATAN PEMILU 2024 MENGGUNAKAN ALGORITMA
NAÏVE BAYES

yang disusun dan diajukan oleh

Muhammad Tamamul Abil Fidaa

20.11.3463

Telah dipertahankan di depan Dewan Penguji
pada tanggal 1 Juli 2025

Susunan Dewan Penguji

Nama Penguji

Rizqi Sukma Kharisma, S.Kom., M.Kom
NIK. 190302215

Novi Prisma Yunita, M.Kom
NIK. 190302526

Mulia Sulistiyono, S.Kom., M.Kom.
NIK. 190302248

Tanda Tangan



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 1 Juli 2025

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusriani, M.Kom.
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Muhammad Tamamul Abil Fidaa

NIM : 20.11.3463

Menyatakan bahwa Skripsi dengan judul berikut:

ANALISIS SENTIMEN DATA TWITTER TERHADAP GUGATAN PEMILU 2024 MENGGUNAKAN ALGORITMA NAÏVE BAYES

Dosen Pembimbing : Mulia Sulistiyono, 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, 1 Juli 2025

Yang Menyatakan,



Muhammad Tamamul Abil Fidaa

HALAMAN PERSEMBAHAN

Puji syukur penulis panjatkan kepada Allah SWT yang telah memberikan nikmat sehat jasmani dan rohani, rahmat, taufik serta hidayah – Nya sehingga penulis dapat menyelesaikan skripsi ini. Shalawat serta salam tak lupa dihanturkan kepada junjungan kita baginda Nabi Muhammad SAW, keluarga, sahabat serta para pengikut beliau hingga akhir zaman. Proses penyelesaian skripsi ini tidak lepas dari bantuan, dukungan, saran, serta kritik dari berbagai pihak. Oleh karena itu, penulis ingin mempersembahkan kepada:

1. Bapak dan Mamah, Umunajmi dan Siti Sunaiyah terima kasih atas doa, semangat, nasehat, dukungan, serta kasih sayang hingga saat ini.
2. Dosen pembimbing Pak Mulia yang sudah membimbing serta memberi masukan dan saran kepada penulis, sehingga saya dapat menyelesaikan skripsi ini.
3. Sahabat Ruli dan Vairuz yang telah memberikan semangat dan pengalaman baik maupun buruk kepada penulis.
4. Kekasih penulis Restu Arya Dhita yang telah memberi segala dukungan, semangat, motivasi, maupun kebersamaan dalam perjalanan ini. Dukungan dan kebaikanmu sangat berarti dan hingga akhirnya skripsi ini dapat terselesaikan.
5. Teman – teman terdekat penulis yang saling berbagi pengalaman, pengetahuan dan memberikan dukungan sehingga skripsi ini dapat terselesaikan yang tidak dapat penulis sebutkan satu persatu.
6. *Last but not least, I wanna thank me. I wanna thank me for believing in me. I wanna thank me for doing all this hard work. I wanna thank me for having no days off. I wanna thank me for never quitting. I wanna thank me for always being a giver and tryna give more than i recieve. I wanna thank me for tryna do more right than wrong. I wanna thank me for just being me at all time.*

Akhir kata, penulis berharap skripsi ini memberikan manfaat bagi pengembangan ilmu serta menjadi kontribusi untuk penelitian selanjutnya.

KATA PENGANTAR

Puji syukur penulis panjatkan kepada Allah SWT atas nikmat dan rahmat-Nya sehingga penulis dapat menyelesaikan skripsi ini sebagai salah satu syarat dalam menyelesaikan pendidikan S1 program studi Informatika di Universitas AMIKOM Yogyakarta. Proses penyelesaian skripsi ini tidak lepas dari bantuan, dukungan, saran, serta kritik dari berbagai pihak, oleh karena itu penulis ingin menyampaikan terima kasih kepada:

1. Allah SWT yang telah memberikan nikmat, rahmat, dan karunia-Nya sehingga penulis dapat menyelesaikan penelitian ini.
2. Kedua orang tua tercinta yang telah memberikan semangat, doa, dukungan, serta kasih sayang tiada henti.
3. Bapak Prof. Dr. M. Suyanto, M.M. selaku Rektor Universitas AMIKOM Yogyakarta.
4. Bapak Mulia Sulistiyono, S.Kom., M.Kom. selaku Dosen Pembimbing yang telah meluangkan waktu dan memberikan bimbingan bantuan, masukan, semangat dan motivasi dalam menyelesaikan skripsi ini.
5. Bapak Rizqi Sukma Kharisma, S.Kom., M.Kom., Ibu Novi Prisma Yunita, M.Kom., Bapak Mulia Sulistiyono, S.Kom., M.Kom. selaku Dewan Penguji atas segala masukan, kritik, dan saran yang telah diberikan sehingga membantu dalam penyempurnaan skripsi ini.
6. Seluruh pihak yang telah mendukung penulisan skripsi ini yang tidak dapat penulis sebutkan satu persatu.

Penulis menyadari masih banyak terdapat kekurangan sehingga peneliti mengharapkan adanya saran dan kritik yang bersifat membangun demi kesempurnaan skripsi ini. Semoga hasil penelitian ini membawa manfaat bagi pengembangan ilmu.

Yogyakarta, 1 Juli 2025

Muhammad Tamamul Abil Fidaa

DAFTAR ISI

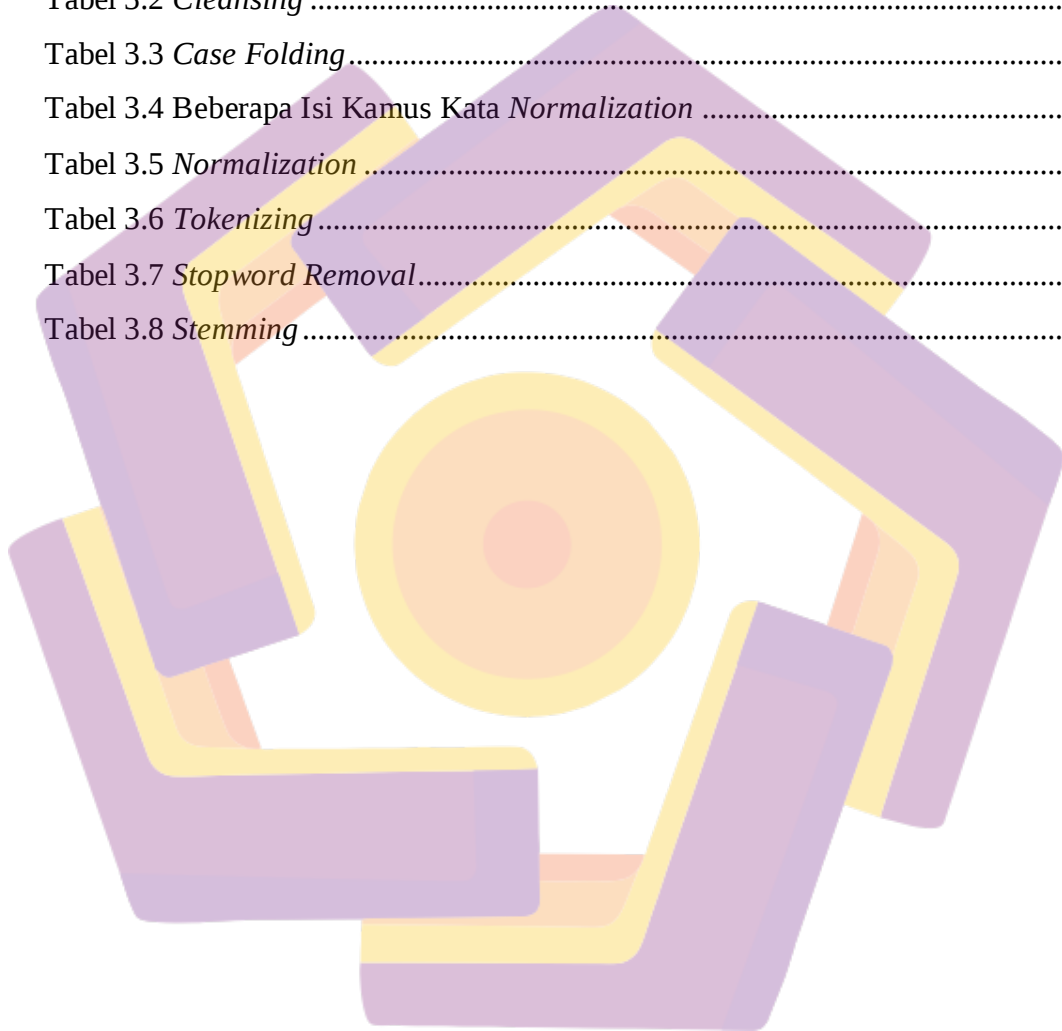
HALAMAN JUDUL	1
HALAMAN PERSETUJUAN	2
HALAMAN PENGESAHAN	3
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	4
HALAMAN PERSEMBAHAN	5
KATA PENGANTAR	6
DAFTAR ISI	7
DAFTAR TABEL	10
DAFTAR GAMBAR	11
DAFTAR LAMBANG DAN SINGKATAN	12
DAFTAR ISTILAH	13
INTISARI	14
ABSTRACT	15
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.6 Sistematika Penulisan	4
BAB II TINJAUAN PUSTAKA.....	5
2.1 Studi Literatur	5
2.2 Dasar Teori	10

2.2.1	<i>Data Mining</i>	10
2.2.2	Analisis Sentimen.....	10
2.2.3	<i>Text Mining</i>	10
2.2.4	Twitter	10
2.2.5	<i>Python</i>	11
2.2.6	Algoritma Naïve Bayes	11
2.2.7	<i>Preprocessing Data</i>	12
2.2.8	SMOTE	12
2.2.9	<i>Confusion Matrix</i>	13
BAB III METODE PENELITIAN.....		15
3.1	Objek Penelitian.....	15
3.2	Alur Penelitian	15
3.2.1	Pengumpulan Data	16
3.2.2	<i>Preprocessing Data</i>	17
3.2.3	<i>Word Cloud</i>	22
3.2.4	SMOTE	22
3.2.5	Klasifikasi Naïve Bayes	23
3.2.6	Evaluasi.....	23
3.3	Alat dan Bahan.....	23
3.3.1	Data Penelitian.....	23
3.3.2	Alat/instrumen	23
BAB IV HASIL DAN PEMBAHASAN		25
4.1	Pengumpulan Data	25
4.2	<i>Preprocessing Data</i>	26
4.2.1	<i>Cleansing</i>	26

4.2.2	<i>Case Folding</i>	27
4.2.3	<i>Normalization</i>	28
4.2.4	<i>Tokenizing</i>	28
4.2.5	<i>Stopword Removal</i>	29
4.2.6	<i>Stemming</i>	30
4.3	Pelabelan Data.....	31
4.4	SMOTE (<i>Synthetic Minority Over-sampling Technique</i>).....	32
4.5	Pembagian Data Latih dan Data Uji	33
4.6	Klasifikasi dengan Algoritma <i>Naïve Bayes</i>	33
4.7	Visualisasi.....	35
BAB V PENUTUP		38
5.1	Kesimpulan.....	38
5.2	Saran	38
REFERENSI.....		40
LAMPIRAN		44

DAFTAR TABEL

Tabel 2.1 Keaslian Penelitian	8
Tabel 2.2 <i>Confusion Matrix</i>	13
Tabel 3.1 Hasil <i>Crawling</i>	16
Tabel 3.2 <i>Cleansing</i>	17
Tabel 3.3 <i>Case Folding</i>	18
Tabel 3.4 Beberapa Isi Kamus Kata <i>Normalization</i>	19
Tabel 3.5 <i>Normalization</i>	19
Tabel 3.6 <i>Tokenizing</i>	20
Tabel 3.7 <i>Stopword Removal</i>	21
Tabel 3.8 <i>Stemming</i>	22

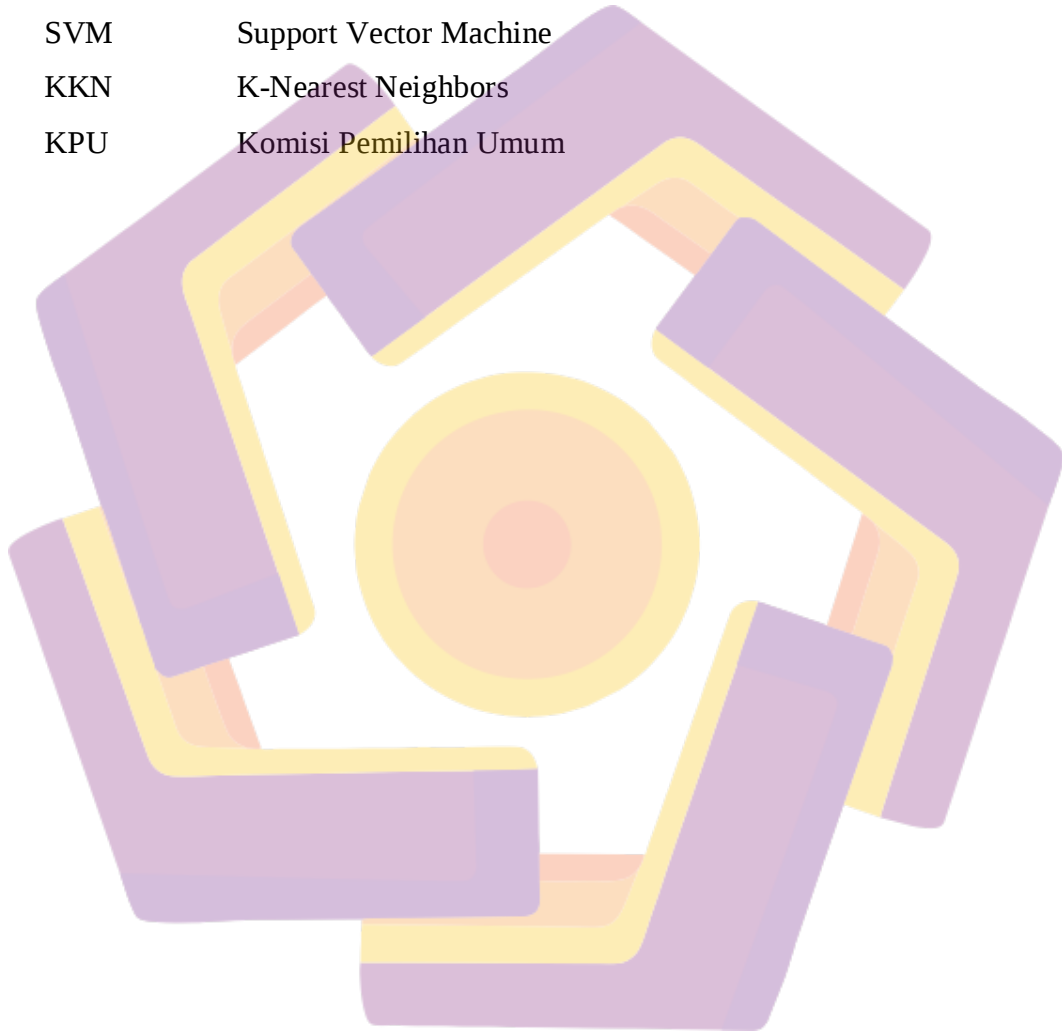


DAFTAR GAMBAR

Gambar 3.1 Alur Penelitian	15
Gambar 4.1 <i>Source Code auth token</i>	25
Gambar 4.2 <i>Source Code Crawling Data</i>	25
Gambar 4.3 Hasil <i>Crawling</i>	25
Gambar 4.4 <i>Source Code Cleansing</i>	26
Gambar 4.5 Hasil <i>Cleansing</i>	27
Gambar 4.6 <i>Source Code Case Folding</i>	27
Gambar 4.7 Hasil <i>Case Folding</i>	27
Gambar 4.8 <i>Source Code Membaca Kamuskatabaku</i>	28
Gambar 4.9 <i>Source Code Normalization</i>	28
Gambar 4.10 Hasil <i>Normalization</i>	28
Gambar 4.11 <i>Source Code Tokenizing</i>	29
Gambar 4.12 Hasil <i>Tokenizing</i>	29
Gambar 4.13 <i>Source Code Stopword Removal</i>	29
Gambar 4.14 Hasil <i>Stopword Removal</i>	30
Gambar 4.15 <i>Source Code Stemming</i>	30
Gambar 4.16 Hasil <i>Stemming</i>	30
Gambar 4.17 <i>Source Code Pelabelan</i>	31
Gambar 4.18 Hasil Pelabelan Data	31
Gambar 4.19 <i>Source Code SMOTE</i>	32
Gambar 4.20 Hasil sebelum dan sesudah SMOTE	33
Gambar 4.21 <i>Source Code Data Latih dan Data Uji</i>	33
Gambar 4.22 <i>Source Code Multinomial Naïve Bayes</i>	34
Gambar 4.23 Hasil <i>Multinomial Naïve Bayes</i>	34
Gambar 4.24 <i>Confusion Matrix</i>	35
Gambar 4.25 <i>Word Cloud Sentimen Positif</i>	36
Gambar 4.26 <i>Word Cloud Sentimen Negatif</i>	37
Gambar 4.27 <i>Word Cloud Sentimen Netral</i>	37

DAFTAR LAMBANG DAN SINGKATAN

csv	Comma-Separated Values
NLTK	Natural Language Toolkit
NLP	Natural Language Processing
URL	Uniform Resource Locator
SVM	Support Vector Machine
KKN	K-Nearest Neighbors
KPU	Komisi Pemilihan Umum



DAFTAR ISTILAH

Crawling	Metode otomatis mengambil data menggunakan program yang disebut web crawler.
Dataset	Kumpulan data yang terorganisir yang digunakan sebagai dasar dalam analisis.
Machine Learning	Cabang dari kecerdasan buatan yang memungkinkan untuk belajar dari data.
Real-time	Proses atau sistem yang memproses data secara langsung.
Hashtag	Untuk menandai topik atau kata kunci tertentu.
Mention	Untuk menyebut atau menandai akun orang lain dalam sebuah postingan atau komentar.
Emoticon	Untuk mengungkapkan emosi atau ekspresi wajah.

INTISARI

Pemilihan Umum (Pemilu) 2024 di Indonesia merupakan peristiwa politik besar yang memunculkan beragam opini dan tanggapan dari masyarakat, terutama melalui media sosial seperti Twitter. Banyak masyarakat menyampaikan pendapatnya terkait gugatan terhadap hasil pemilu yang diajukan ke Mahkamah Konstitusi. Fenomena ini menunjukkan perlunya analisis opini publik secara sistematis agar dapat diketahui sentimen dominan masyarakat terhadap isu tersebut.

Penelitian ini bertujuan untuk menganalisis sentimen tweet publik terkait gugatan Pemilu 2024 menggunakan algoritma *Naïve Bayes*. Data tweet berbahasa Indonesia dikumpulkan melalui metode *crawling* menggunakan *tweet harvest*. Setelah dilakukan tahapan *preprocessing* seperti *cleansing*, *case folding*, normalisasi, *tokenizing*, *stopword*, dan *stemming*, data diberi label sentimen secara otomatis menggunakan kamus Inset Lexicon. Untuk mengatasi ketidakseimbangan jumlah data antar kelas sentimen, digunakan metode SMOTE (*Synthetic Minority Oversampling Technique*). Proses klasifikasi dilakukan menggunakan algoritma Multinomial *Naïve Bayes* dan dievaluasi menggunakan *confusion matrix*.

Hasil evaluasi menunjukkan bahwa model memiliki performa yang baik, dengan nilai akurasi sebesar 80,64%, *precision* 81,91%, *recall* 78,96%, dan F1-score sebesar 79,38%. Selain itu, visualisasi *word cloud* digunakan untuk menggambarkan kata-kata yang sering muncul dalam setiap kategori sentimen. Hasil penelitian ini menunjukkan bahwa algoritma *Naïve Bayes* cukup efektif dalam mengklasifikasikan sentimen terkait gugatan Pemilu. Dengan demikian, penelitian ini diharapkan memberikan kontribusi dalam pemahaman opini publik serta menjadi referensi bagi penelitian selanjutnya.

Kata kunci: Analisis Sentimen, *Naïve Bayes*, Twitter, Gugatan Pemilu, SMOTE.

ABSTRACT

The 2024 General Election in Indonesia is a major political event that has given rise to various opinions and responses from the public, especially through social media such as Twitter. Many people have expressed their opinions regarding the lawsuit against the election results filed with the Constitutional Court. This phenomenon shows the need for a systematic analysis of public opinion in order to determine the dominant sentiment of the public on the issue.

This study aims to analyze public tweet sentiment related to the 2024 Election lawsuit using the Naïve Bayes algorithm. Indonesian language tweet data were collected through the crawling method using tweet harvest. After carrying out preprocessing stages such as cleansing, case folding, normalization, tokenizing, stopwords, and stemming, the data was automatically labeled with sentiment using the Inset Lexicon dictionary. To overcome the imbalance in the amount of data between sentiment classes, the SMOTE (Synthetic Minority Oversampling Technique) method was used. The classification process was carried out using the Multinomial Naïve Bayes algorithm and evaluated using a confusion matrix.

The evaluation results show that the model has good performance, with an accuracy value of 80.64%, precision of 81.91%, recall of 78.96%, and F1-score of 79.38%. In addition, word cloud visualization is used to illustrate words that often appear in each sentiment category. The results of this study indicate that the Naïve Bayes algorithm is quite effective in classifying sentiments related to election lawsuits. Thus, this study is expected to contribute to the understanding of public opinion and become a reference for further research.

Keyword: Sentiment Analysis, Naïve Bayes, Twitter, Election Lawsuits, SMOTE