

**ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP ISU  
LEGALISASI GANJA DI INDONESIA MENGGUNAKAN  
METODE NAÏVE BAYES**

**SKRIPSI**

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



disusun oleh

**MUHAMMAD FAHMI ATTAQWA**

**19.11.2597**

Kepada

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

**2023**

**ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP ISU  
LEGALISASI GANJA DI INDONESIA MENGGUNAKAN  
METODE NAÏVE BAYES**

**SKRIPSI**

untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi Informatika



disusun oleh

**MUHAMMAD FAHMI ATTAQWA**

**19.11.2597**

Kepada

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

**2023**

**HALAMAN PERSETUJUAN**

**SKRIPSI**

**ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP ISU  
LEGALISASI GANJA DI INDONESIA MENGGUNAKAN METODE  
NAÏVE BAYES**

yang disusun dan diajukan oleh

**MUHAMMAD FAHMI ATTAQWA**

**19.11.2597**

telah disetujui oleh Dosen Pembimbing Skripsi  
pada tanggal 23 Juni 2023

Dosen Pembimbing,



**Anna Balta M.Kom**

**NIK. 190302290**

**HALAMAN PENGESAHAN**

**SKRIPSI**

**ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP ISU  
LEGALISASI GANJA DI INDONESIA MENGGUNAKAN MEOTDE  
NAÏVE BAYES**

yang disusun dan diajukan oleh

**MUHAMMAD FAHMI ATTAQWA**

19.11.2597

Telah dipertahankan di depan Dewan Penguji  
pada tanggal 23 Juni 2023

**Susunan Dewan Penguji**

**Nama Penguji**

**Tanda Tangan**

Anna Baita, M.Kom  
NIK. 190302290

Wiwi Widayani, M.Kom  
NIK. 190302272

Ali Mustopa, M.Kom  
NIK. 190302192



Skripsi ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Sarjana Komputer  
Tanggal 23 Juni 2023

**DEKAN FAKULTAS ILMU KOMPUTER**



Hanif Al Fatta, S.Kom., M.Kom.  
NIK. 190302096

## HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : MUHAMMAD FAHMI ATTAQWA  
NIM : 19.11.2597

Menyatakan bahwa Skripsi dengan judul berikut:

**ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP ISU  
LEGALISASI GANJA DI INDONESIA MENGGUNAKAN METODE  
NAÏVE BAYES**

Dosen Pembimbing: Anna Baita, 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, 23 Juni 2023

Yang Menyatakan,



Muhammad Fahmi Attaqwa

## HALAMAN PERSEMBAHAN

Alhamdulillahirabbil'alamin, segala puji bagi ALLAH SWT berkat rahmat serta hidayah-Nya sehingga penyusunan skripsi ini dapat terselesaikan dengan lancar. Skripsi ini saya persembahkan untuk:

1. Orang tua saya, dengan penuh cinta dan kesetiaan, selalu mendoakan, memberikan restu, dan memberikan dukungan motivasi kepada saya.
2. Ibu Anna Baita, M.Kom, sebagai dosen pembimbing saya, selalu memberikan bimbingan dengan kesabaran dan ikhlas, membantu saya ketika saya menghadapi kesulitan dalam menyusun skripsi.
3. Teman-teman sekelas IF-01, terutama Takeshi Castle, telah menjadi sahabat setia selama hampir empat tahun. Mereka selalu memberikan motivasi dan dukungan yang tak ternilai bagi saya dan menjadi teman terbaik yang saya miliki.
4. Semua teman yang mengenal saya, meskipun tidak dapat saya sebutkan satu per satu. Dukungan dan motivasi yang telah kalian berikan kepada saya sangat berarti.

## KATA PENGANTAR

Dengan segala puji dan syukur kepada Allah SWT yang telah melimpahkan rahmat-Nya, petunjuk-Nya, dan karunia-Nya, saya berhasil menyelesaikan penyusunan skripsi dengan judul "Analisis Sentimen Pengguna Twitter Terhadap Isu legalisasi Ganja di Indonesia Menggunakan Metode Naïve Bayes". Skripsi ini diajukan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana di jurusan Informatika, Universitas Amikom Yogyakarta. Selama proses penyusunannya, saya mendapatkan banyak bantuan dari berbagai pihak, dan dengan ini saya ingin menyampaikan rasa terima kasih kepada:

1. Ibu Anna Baita M, Kom. Sebagai dosen pembimbing yang telah memberikan arahan, pengarahan, dan dukungan yang sangat berarti dalam penulisan skripsi ini.
2. Kedua orang tua saya yang telah selalu memberikan doa, restu serta dukungan motivasi kepada saya.
3. Seluruh dosen yang telah memberikan ilmu, bimbingan dan pengalaman yang sangat berharga bagi penulis selama masa perkuliahan.
4. Semua pihak yang telah membantu dan memberikan dukungan dan doa dalam penyelesaian skripsi ini.

Penulis menyadari bahwa skripsi ini masih memiliki kekurangan. Penulis mengharapkan kritik dan saran yang membangun untuk peningkatan laporan selanjutnya. Terima kasih atas segala bantuan dan dukungan. Wassalamu'alaikum  
Wr. Wb.

Yogyakarta, 25 Juni 2023



Muhammad Fahmi Attaqwa

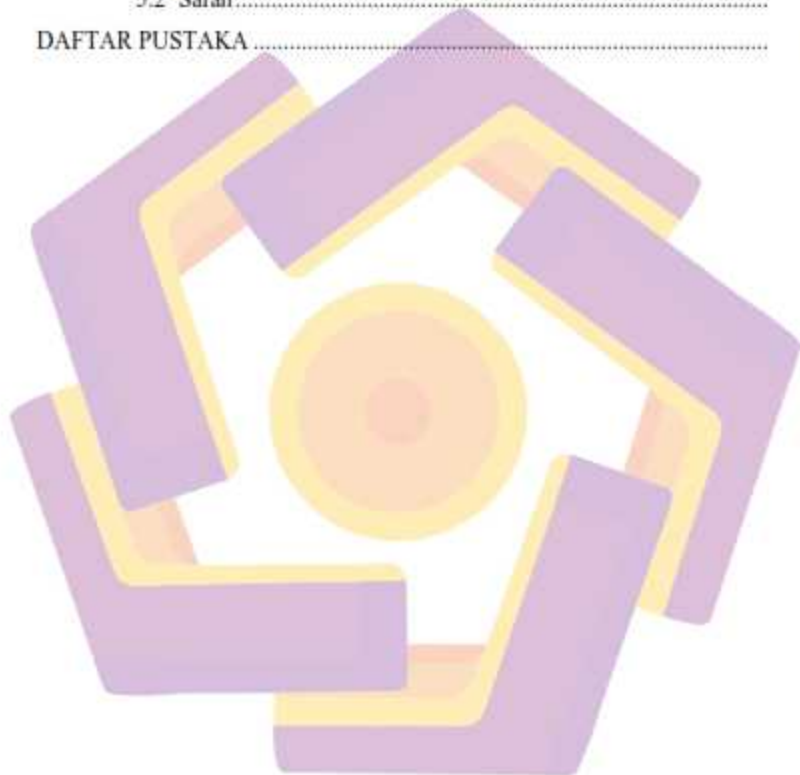
## DAFTAR ISI

HALAMAN JUDUL .....	i
HALAMAN PERSETUJUAN .....	ii
HALAMAN PENGESAHAN .....	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI .....	iv
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
<b>BAB I PENDAHULUAN .....</b>	<b>1</b>
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	2
1.3 Batasan Masalah .....	3
1.4 Tujuan Penelitian .....	3
1.5 Manfaat Penelitian .....	3
1.5.1 Manfaat Teoritis .....	3
1.5.2 Manfaat Praktis .....	3
1.6 Sistematika Penulisan .....	4
<b>BAB II TINJAUAN PUSTAKA .....</b>	<b>5</b>
2.1 Studi Literatur .....	5
2.2 Dasar Teori .....	12
2.2.1 Legalisasi Ganja .....	12
2.2.2 Twitter .....	12
2.2.3 Python .....	13
2.2.4 Google Colaboratory .....	13



2.2.5	Klasifikasi.....	13
2.2.6	Analisis Sentimen.....	14
2.2.7	TF-IDF.....	14
2.2.8	Naïve Bayes.....	15
2.2.9	Confusion Matrix .....	16
<b>BAB III</b>	<b>METODE PENELITIAN .....</b>	<b>17</b>
3.1	Objek Penelitian.....	17
3.2	Alur Penelitian.....	17
3.2.1	Crawling Data.....	18
3.2.2	Data Preprocessing .....	19
3.2.2	Translate Data.....	24
3.2.3	Labeling Data.....	26
3.2.4	Term Weighting .....	27
3.2.5	Splitting Data.....	31
3.2.6	Modeling Naïve Bayes .....	31
3.2.7	Evaluation.....	32
3.3	Alat dan Bahan .....	32
3.3.1	Bahan Penelitian.....	32
3.3.2	Alat.....	32
<b>BAB IV</b>	<b>HASIL DAN PEMBAHASAN .....</b>	<b>33</b>
4.1	Pengumpulan Data dan Pengolahan Data .....	33
4.2	Evaluation .....	34
4.2.1	Pengujian 70:30.....	34
4.2.2	Pengujian 80:20.....	36
4.2.3	Pengujian 90:10.....	37
4.2.4	Pengujian Unigram.....	38
4.2.5	Pengujian Bigram .....	40
4.2.6	Pengujian Trigram.....	41
4.3	Visualisasi Data.....	42
4.3.1	Unigram, Bigram dan Trigram Sentimen Positif.....	42
4.3.2	Unigram, Bigram dan Trigram Sentimen Negatif.....	44

4.4 Diskusi Hasil .....	45
4.4.1 Evaluation.....	45
4.4.2 Visualisasi Data.....	46
BAB V PENUTUP .....	48
5.1 Kesimpulan.....	48
5.2 Saran.....	48
DAFTAR PUSTAKA .....	49



## DAFTAR TABEL

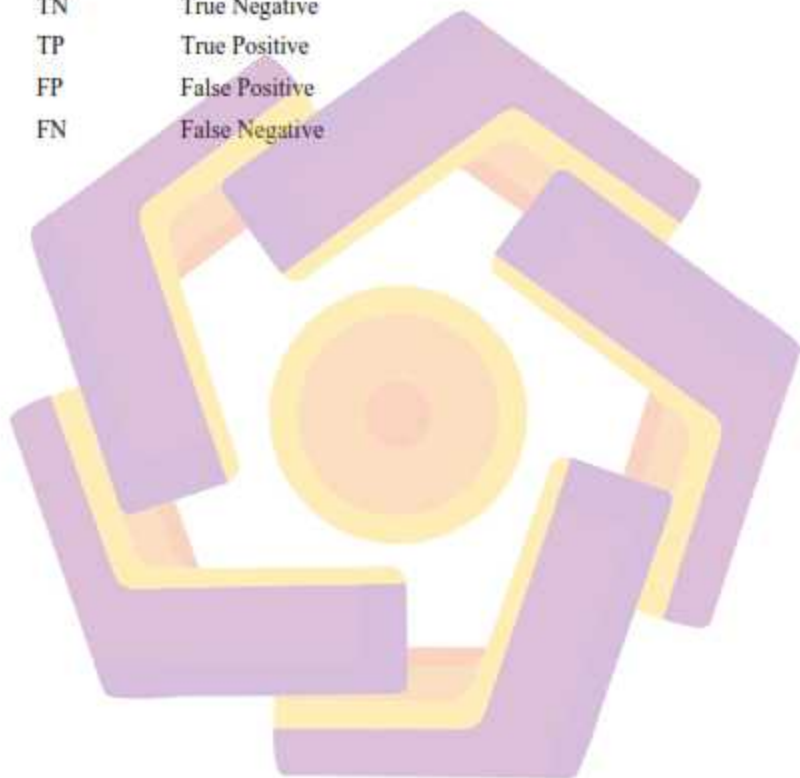
Tabel 2.1 Keaslian Penelitian .....	11
Tabel 3. 1 Hasil <i>Case Folding</i> .....	21
Tabel 3. 2 Hasil Normalisasi.....	22
Tabel 3. 3 Hasil <i>Stopword</i> .....	23
Tabel 3. 4 Hasil <i>Stemming</i> .....	24
Tabel 3. 5 Hasil <i>Translate data</i> .....	26
Tabel 3. 6 Hasil <i>Text Preprocessing</i> .....	27
Tabel 3. 7 <i>Term Frequency (TF)</i> dan <i>Document Frequency (DF)</i> .....	29
Tabel 3. 8 Perhitungan <i>Inverse Document Frequency (IDF)</i> .....	30
Tabel 3. 9 Perhitungan bobot tiap kata .....	31
Tabel 4. 1 Total Penghapusan Data .....	34
Tabel 4. 2 Perbedaan Jumlah Sentimen Hasil dan Asli 70:30 .....	35
Tabel 4.3 Perbedaan Jumlah Sentimen Hasil dan Asli 80:20 .....	37
Tabel 4.4 Perbedaan Jumlah Sentimen Hasil dan Asli 90:10 .....	38
Tabel 4. 5 Perbedaan Jumlah Sentimen Hasil dan Asli Bigram .....	41
Tabel 4. 6 Perbedaan Jumlah Sentimen Hasil dan Asli Trigram .....	42
Tabel 4. 7 Hasil Visualisasi Data .....	47

## DAFTAR GAMBAR

Gambar 3.1 Alur Penelitian .....	17
Gambar 3.2 Hasil Pengumpulan Data yang Sudah Digabungkan .....	18
Gambar 3.3 Jumlah Label Positif dan Negatif.....	27
Gambar 4.1 Perbandingan Sentimen Positif dan Negatif .....	34
Gambar 4.2 Hasil Pengujian Data <i>Trainig</i> 70% dan <i>Testing</i> 30% .....	35
Gambar 4.3 Hasil <i>Confusion Matrix</i> 70% dan 30% .....	35
Gambar 4.4 Hasil Pengujian Data <i>Trainig</i> 80% dan <i>Testing</i> 20% .....	36
Gambar 4.5 Hasil <i>Confusion Matrix</i> 80% dan 20% .....	36
Gambar 4.6 Hasil Pengujian Data <i>Trainig</i> 90% dan <i>Testing</i> 10%.....	37
Gambar 4.7 Hasil <i>Confusion Matrix</i> 90% dan 10% .....	38
Gambar 4.8 Hasil Pengujian Unigram .....	39
Gambar 4.9 Hasil <i>Confusion Matrix Unigram</i> .....	39
Gambar 4.10 Hasil Pengujian Bigram .....	40
Gambar 4.11 <i>Confusion Matrix Bigram</i> .....	40
Gambar 4.12 Hasil Pengujian Trigram .....	41
Gambar 4.13 Hasil <i>Confusion Matrix Trigram</i> .....	42
Gambar 4.14 Unigram, Bigram dan Trigram pada sentiment positif.....	43
Gambar 4.15 Unigram, Bigram dan Trigram Sentimen Negatif .....	44
Gambar 4.16 Hasil Pengujian 70:30, 80:20 dan 90:10 .....	45
Gambar 4.17 Hasil Pengujian N-Gram.....	46

## DAFTAR LAMBANG DAN SINGKATAN

THC	Tetrahidrokanabinol
TF	Term Frequency
IDF	Inverse Documen Frequency
TN	True Negative
TP	True Positive
FP	False Positive
FN	False Negative



## DAFTAR ISTILAH

Analgesic	obat pereda rasa nyeri
Euphoria	perasaan nyaman dan bahagia secara berlebihan
Glaucoma	kondisi kelainan medis karena kerusakan saraf mata
Tetrahidrokanabinol	kanabinoid utama di dalam tumbuhan ganja
Isu	hal hal yang belum tentu kebenarannya
Legalisasi	proses mengubah hukum dari illegal menjadi legal
Crawling	proses otomatis dalam pengumpulan data
Preprocessing	pengolahan data atau pemrosesan data
Sentiment	pendapat atau pandangan



## INTISARI

Isu legalisasi ganja menimbulkan berbagai respon dari masyarakat semenjak terjadinya penggunaan ganja untuk medis seperti yang terjadi pada fidelis yang menggunakan ganja untuk pengobatan istrinya dan ibu-ibu yang membawa poster bertuliskan legalkan ganja medis yang akan digunakan untuk pengobatan anaknya. Respon masyarakat biasanya ditulis di media sosial. Salah satu media sosial yang digunakan untuk menyampaikan aspirasi dan opini adalah twitter. Oleh karena itu peneliti tertarik melakukan penelitian analisis sentiment dalam topik isu legalisasi ganja dan menggunakan data dari sosial media twitter.

Penelitian ini menggunakan metode naïve bayes dengan jumlah data 2133 tweet. Hasil didapatkan jumlah sentiment dari klasifikasi menggunakan naïve bayes berjumlah 1269 untuk sentiment positif dan 864 untuk sentiment negatif. Performa naïve bayes yang paling bagus dihasilkan dengan penggunaan unigram pada pembobotan tf-idf dan menggunakan data training 90% dengan hasil akurasi sebesar 78, presisi 73, *recall* 96 dan *f1 score* 83.

**Kata kunci:** naïve bayes, legalisasi ganja, analisis sentimen, twitter, n-gram.

## **ABSTRACT**

*The issue of cannabis legalization has elicited various responses from the public since the use of medical cannabis occurred, as happened to Fidelis who used cannabis for the treatment of his wife and mothers who carried posters that said legalize medical cannabis to be used for the treatment of their children. Community responses are usually written on social media. One of the social media used to convey aspirations and opinions is Twitter. Therefore researchers are interested in conducting sentiment analysis research on the topic of cannabis legalization issues and using data from social media Twitter.*

*This study uses the Naïve Bayes method with a total of 2133 tweets. The results showed that the number of sentiments from the classification using naïve Bayes was 1269 for positive sentiment and 864 for negative sentiment. The best performance of Naïve Bayes is produced by using the unigram on tf-idf weighting and using 90% training data with an accuracy of 78, precision 73, recall 96 and f1 score 83.*

**Keyword:** *naïve bayes, legalization of cannabis, sentiment analysis, twitter, n-gram.*