

**SENTIMENT ANALISIS PENGGUNA TWITTER TERHADAP
KEBIJAKAN PEMERINTAH MENGENAI HARGA MINYAK
GORENG DI INDONESIA MENGGUNAKAN METODE NAIVE
BAYES CLASSIFIER**

SKRIPSI

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



disusun oleh

Tegar Sukma Abdi

17.11.1729

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2023**

**SENTIMENT ANALISIS PENGGUNA TWITTER TERHADAP
KEBIJAKAN PEMERINTAH MENGENAI HARGA MINYAK
GORENG DI INDONESIA MENGGUNAKAN METODE
NAIVE BAYES CLASSIFIER**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

Tegar Sukma Abdi

17.11.1729

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2023

HALAMAN PERSETUJUAN

SKRIPSI

**SENTIMENT ANALISIS PENGGUNA TWITTER TERHADAP
KEBIJAKAN PEMERINTAH MENGENAI HARGA MINYAK
GORENG DI INDONESIA MENGGUNAKAN METODE
NAIVE BAYES CLASSIFER**

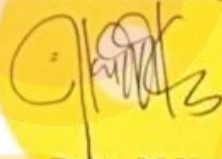
yang disusun dan diajukan oleh

Tegar Sukma Abdi

17.11.1729

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 22 Agustus 2023

Dosen Pembimbing,



Anna Baita, M.Kom

NIK. 190302290

**HALAMAN PENGESAHAN
SKRIPSI**

**SENTIMENT ANALISIS PENGGUNA TWITTER TERHADAP
KEBIJAKAN PEMERINTAH MENGENAI HARGA MINYAK
GORENG DI INDONESIA MENGGUNAKAN METODE
NAIVE BAYES CLASSIFER**

yang disusun dan diajukan oleh

Tegar Sukma Abdi

17.11.1729

Telah dipertahankan di depan Dewan Penguji
pada tanggal 22 Agustus 2023

Susunan Dewan Penguji

Nama Penguji

Anna Baita, M.Kom
NIK. 190302290

Theopilus Bayu Sasongko, S.Kom, M.Eng
NIK. 190302375

Arif Akbarul Huda, S.Si, M.Eng
NIK. 190302287

Tanda Tangan



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 22 Agustus 2023

DEKAN FAKULTAS ILMU KOMPUTER



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

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertanda tangan di bawah ini,

Nama mahasiswa : **TEGAR SUKMA ABDI**

NIM : **17.11.1729**

Menyatakan bahwa Skripsi dengan judul berikut:

**SENTIMENT ANALISIS PENGGUNA TWITTER TERHADAP
KEBIJAKAN PEMERINTAH MENGENAI HARGA MINYAK GORENG
DI INDONESIA MENGGUNAKAN METODE NAIVE 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,
Yang Menyatakan.



HALAMAN PERSEMBAHAN

Puji dan Syukur atas kehadiran Allah SWT, yang telah melimpahkan rahmat, hidayah dan karunia-Nya, atas kemudahan, petunjuk dan kekuatan yang diberikan-Nya, sehingga penulis dapat menyelesaikan penyusunan skripsi yang berjudul

“Analisis Sentimen Pengguna Twitter Terhadap Kebijakan Pemerintah Mengenai Harga Minyak Goreng di Indonesia Menggunakan Metode Naive Bayes Classifier”,

dan diajukan sebagai salah satu syarat untuk mendapat gelar Sarjana pada jurusan Informatika, Universitas Amikom Yogyakarta. Karya tulis ini penulis persembahkan kepada :

1. Pertama saya ucapkan terima kasih kepada ayah saya Muso Susanto, yang sudah menyiapkan kebutuhan finansial saya dalam menempuh pendidikan hingga berhasil mencapai titik ini sebagai sarjana pertama dalam keluarga saya.
2. Kedua saya ucapkan terima kasih kepada ibu saya Eka Sri Indrasasti, yang selalu mendukung dan menyemangati saya dalam menjalani pendidikan perguruan tinggi hingga merantau ke luar daerah. Kemudian saya ucapkan juga terima kasih kepada ibu saya yang selalu mengingatkan saya untuk beribadah dan selalu memberikan doa yang terbaik untuk saya.
3. Ketiga saya ucapkan terima kasih kepada adik saya Novia Kartika Santi yang selalu mendukung dan menyemangati saya dalam menempuh pendidikan di luar daerah.
4. Terakhir, saya ucapkan terima kasih kepada teman, saudara, dan orang yang selalu ada untuk membantu dan menyemangati saya dalam penulisan skripsi ini

KATA PENGANTAR

KATA PENGANTAR

Puji dan Syukur atas kehadiran Allah SWT yang telah melimpahkan rahmat,

hidayah dan karunia-Nya, atas kemudahan, petunjuk dan kekuatan yang diberikan-Nya, sehingga penulis dapat menyelesaikan penyusunan skripsi yang berjudul

“Analisis Sentimen Pengguna Twitter Terhadap Kebijakan Pemerintah Mengenai Harga Minyak Goreng di Indonesia Menggunakan Metode Naive Bayes Classifier”, dan diajukan sebagai salah satu syarat untuk mendapat gelar Sarjana pada jurusan Informatika, Universitas Amikom Yogyakarta. Dalam penyusunannya, penulis memperoleh banyak bantuan dari berbagai pihak, oleh karena itu penulis mengucapkan terimakasih yang sebesar-besarnya kepada :

1. Ibu Anna Baita, M. Kom sebagai dosen pembimbing yang telah memberikan kontribusi berupa bimbingan, motivasi, kritik dan saran dalam menyelesaikan skripsi ini.
2. Semua Dosen Fakultas Ilmu Komputer Universitas Amikom Yogyakarta yang tidak bisa penulis sebutkan satu persatu. Terimakasih atas ilmu yang sudah diajarkan selama ini.
3. Kedua orang tua dan keluarga besar yang senantiasa mendoakan dan menyemangati serta memberi dukungan.
4. Semua pihak yang telah membantu dalam penyusunan skripsi ini dari awal hingga akhir.

Tidak ada kata selain terima kasih yang sebesar-besarnya kepada semua pihak yang telah membantu saya sampai titik ini. Penulis menyadari bahwa penulisan skripsi ini masih memiliki banyak kekurangan. Untuk semua kritik dan saran yang bersifat membangun sangat penulis harapkan dan akan diterima dengan lapang dada.

Yogyakarta, 15 September 2023



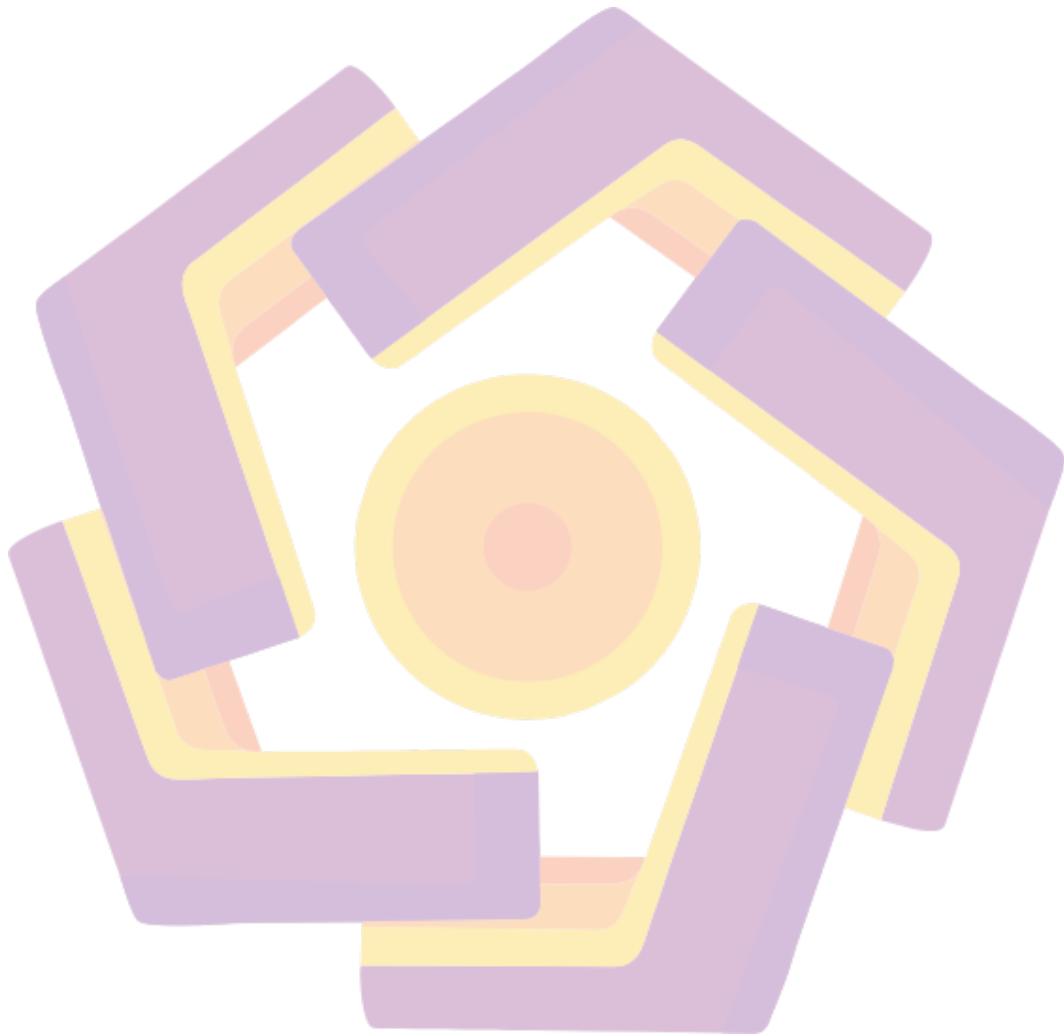
Tegar Sukma Abdi

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
INTISARI.....	xii
ABSTRACT.....	xiii
BAB I Pendahuluan	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah.....	3
1.4 Maksud dan Tujuan Penelitian.....	3
1.5 Manfaat Penelitian.....	3
1.6 Metode Penelitian.....	4
1.7 Sistematika Penulisan.....	4
BAB II TINJAUAN PUSTAKA.....	6
2.1 Studi Literatur.....	6
2.2 Dasar Teori.....	13
2.2.1 Twitter.....	13

2.2.2	Analisis Sentimen	13
2.2.3	Text <i>Pre-Processing</i>	14
2.2.4	Pembobotan TF-IDF	15
2.2.5	Naïve bayes classifier.....	16
2.2.6	Confusion Matrix	18
BAB III METODE PENELITIAN.....		20
3.1	Objek Penelitian	20
3.2	Alur penelitian.....	20
3.2.1	Observasi.....	21
3.2.2	Studi Pustaka.....	22
3.2.3	Pengumpulan Data	22
3.2.4	<i>Pre-Processing</i> Data	24
3.2.5	Pembobotan dengan <i>TF-IDF</i>	29
3.2.6	Uji Coba	30
3.2.7	Hasil dan Kesimpulan	30
3.3	Alat dan Bahan Penelitian	30
3.3.1	Alat.....	30
3.3.2	Bahan Penelitian.....	31
BAB IV HASIL DAN PEMBAHASAN		32
4.1	Pengumpulan dan Pengolahan Dataset.....	32
4.2	Pembobotan TF-IDF.....	33
4.3	Visualisasi Kata	34
4.3.1	Visualisasi Kata Positif	34
4.3.2	Visualisasi Kata Negatif.....	35
4.4	Implementasi Naïve Bayes	36

4.5 Analisis Hasil Pengujian	41
BAB V PENUTUP.....	41
DAFTAR PUSTAKA	42



DAFTAR TABEL

Tabel 2.1 Tabel Perbandingan Kajian Hasil Penelitian	8
Tabel 2.2 <i>Confusion Matrix</i>	18
Tabel 3.1 <i>Output Text Cleaning</i>	24
Tabel 3.2 <i>Output Proses Stopword Removal</i>	26
Tabel 3.3 <i>Output Labeling</i>	27
Tabel 3.4 <i>Output Proses Stemming</i>	29
Tabel 4.1 Hasil Pembobotan TF-IDF	33
Tabel 4.1 10 Kata yang Sering Muncul pada Sentimen Positif	35
Tabel 4.2 10 Kata yang Sering Muncul pada Sentimen Negatif	36
Tabel 4.3 <i>Dataset Split</i>	36
Tabel 4.4 <i>Confusion Matrix</i> Perbandingan A	37
Tabel 4.5 <i>Confusion Matrix</i> Perbandingan B	38
Tabel 4.6 <i>Confusion Matrix</i> Perbandingan C	39
Tabel 4.7 <i>Confusion Matrix</i> Perbandingan D	40
Tabel 4.8 <i>Confusion Matrix</i> Perbandingan E	40

DAFTAR GAMBAR

Gambar 3.1 Alur Penelitian	21
Gambar 3.2 Observasi Menggunakan Kata Kunci	22
Gambar 3.3 fitur pengambilan data pada website exportcomments.com	23
Gambar 3.4 Data Tweets Kebijakan Pemerintah Harga Minyak Goreng	23
Gambar 3.5 <i>Text Cleaning</i>	24
Gambar 3.6 <i>Stopword Removal</i>	25
Gambar 3.7 Menghapus Duplikat	27
Gambar 3.8 <i>Tokenizing</i>	28
Gambar 3.9 <i>Stemming</i>	29
Gambar 4.1 Pembagian Dataset Sentiment	33
Gambar 4.2 Visualisasi Sentimen yang Sering Muncul	34
Gambar 4.3 Sentimen Positif yang Sering Muncul	34
Gambar 4.4 Sentimen Negatif yang Sering Muncul	35
Gambar 4.5 Performa Naïve Bayes Persamaan A Label Otomatis	37
Gambar 4.6 Performa Naïve Bayes Persamaan A Label Manual	38
Gambar 4.7 Performa Naïve Bayes Persamaan B Label Otomatis	38
Gambar 4.8 Performa Naïve Bayes Persamaan B Label Manual	39
Gambar 4.9 Performa Naïve Bayes Persamaan C Label Otomatis	39
Gambar 4.10 Performa Naïve Bayes Persamaan C Label Manual	39
Gambar 4.11 Performa Naïve Bayes Persamaan D Label Otomatis	40
Gambar 4.12 Performa Naïve Bayes Persamaan D Label Manual	40
Gambar 4.13 Performa Naïve Bayes Persamaan E Label Otomatis.	41
Gambar 4.14 Performa Naïve Bayes Persamaan E Label Manual.	41
Gambar 4.15 Perbandingan Performa Naïve Bayes Label Otomatis	42
Gambar 4.16 Perbandingan Performa Naïve Bayes Label Manual	42

INTISARI

Perkembangan teknologi internet yang begitu pesat, membuat berbagai macam aktivitas manusia bergantung pada kemajuan teknologi salah satunya media sosial. Media sosial, termasuk Twitter, telah menjadi platform yang memainkan peran penting dalam memfasilitasi komunikasi dan berbagi informasi di antara masyarakat global. Dalam konteks ini, Twitter telah menjadi ruang publik virtual di mana pengguna dapat menyampaikan opini, pandangan pribadi, serta mengungkapkan sentimen mereka mengenai berbagai hal, termasuk topik sensitif seperti kelangkaan minyak goreng.

Untuk mengidentifikasi opini masyarakat, dilakukan analisis sentiment yang bertujuan untuk mengetahui apakah opini terhadap minyak goreng bersifat positif atau negatif. Untuk menjalankan analisis sentimen ini, diperlukan algoritma yang mumpuni. Salah satu algoritma yang diterapkan untuk mengklasifikasi data tweet terkait isu kekurangan minyak goreng adalah algoritma Naïve Bayes Classifier (NBC).

Hasil penelitian ini mencakup analisis sentimen tentang kelangkaan minyak goreng, diketahui bahwa sebagian besar pengguna mengekspresikan sentimen negatif terhadap situasi ini, dengan persentase sebesar 59,84% (681 cuitan). Sedangkan pengguna yang menunjukkan sentimen positif sebanyak 40,16% (457 cuitan). Sementara itu, dalam uji coba algoritma Naïve Bayes yang dijalankan sebanyak 5 kali, ditemukan hasil terbaik pada skema pembagian data, dengan 80% digunakan sebagai data latihan (*training*) dan 20% sebagai data pengujian (*testing*). Dari hasil ini, akurasi algoritma mencapai 82%, dengan nilai presisi mencapai 84%, nilai recall sebesar 82%, dan nilai f1-score sebesar 83%.

Kata Kunci: Analisis sentimen, kelangkaan minyak goreng, naïve bayes classifier, twitter.

ABSTRACT

The rapid development of internet technology has made various kinds of human activities depend on technological advances, one of which is social media. Social media, including Twitter, has become a platform that plays an important role in facilitating communication and information sharing among global communities. In this context, Twitter has become a virtual public space where users can express opinions, personal views, and express their sentiments on various matters, including sensitive topics such as the scarcity of cooking oil.

To identify public opinion, a sentiment analysis is carried out which aims to find out whether the opinion on cooking oil is positive or negative. To run this sentiment analysis, a qualified algorithm is needed. One of the algorithms applied to classify tweet data related to the issue of shortage of cooking oil is the Naïve Bayes Classifier (NBC) algorithm.

The results of this study include sentiment analysis about the scarcity of cooking oil, it is known that most users express negative sentiments towards this situation, with a percentage of 59.84% (681 tweets). Meanwhile, users showing positive sentiment were 40.16% (457 tweets). Meanwhile, in testing the Naïve Bayes algorithm which was run 5 times, the best results were found in the data sharing scheme, with 80% used as training data and 20% as testing data. From these results, the accuracy of the algorithm reaches 82%, with a precision value of 84%, a recall value of 82%, and an f1-score value of 83%.

Keywords: *Sentiment analysis, scarcity of cooking oil, naïve Bayes classifier, twitter.*