

**ANALISIS SENTIMEN TWITTER TENTANG KENAIKAN  
BAHAN BAKAR MINYAK DENGAN METODE  
NAIVE BAYES CLASIFIER (NBC)**

**SKRIPSI**

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



disusun oleh

**MOHAMMAD IRHAM**

**19.11.3162**

Kepada

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

**2022**

**ANALISIS SENTIMEN TWITTER TENTANG KENAIKAN  
BAHAN BAKAR MINYAK DENGAN METODE  
NAIVE BAYES CLASIFIER (NBC)**

**SKRIPSI**

untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi Informatika



disusun oleh

**MOHAMMAD IRHAM**

**19.11.3162**

Kepada

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

**2022**

**HALAMAN PERSETUJUAN**

**SKRIPSI**

**ANALISIS SENTIMEN TWITTER TENTANG KENAIKAN  
BAHAN BAKAR MINYAK DENGAN METODE  
NAIVE BAYES CLASIFIER (NBC)**

yang disusun dan diajukan oleh

**MOHAMMAD IRHAM**

**19.11.3162**

telah disetujui oleh Dosen Pembimbing Skripsi  
pada tanggal 24 juni 2023

**Dosen Pembimbing,**



**ERNI SENIWATI, S.KOM, M.Cs**

**NIK. 190302231**

**HALAMAN PENGESAHAN**

**SKRIPSI**

**ANALISIS SENTIMEN TWITTER TENTANG KENAIKAN  
BAHAN BAKAR MINYAK DENGAN METODE  
NAIVE BAYES CLASIFIER (NBC)**

yang disusun dan diajukan oleh

**Mohammad Irham**

**19.11.3162**

Telah dipertahankan di depan Dewan Penguji  
pada tanggal 24 juni 2023

**Susunan Dewan Penguji**

**Nama Penguji**

**Erni Seniwati, S.Kom, M.Cs**  
**NIK. 190302231**

**Rifda Faticha Alfa Aziza, S.Kom, M.Kom**  
**NIK. 190302392**

**Jeki Kuswanto, M.Kom**  
**NIK. 190302456**

**Tanda Tangan**



Skripsi ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Sarjana Komputer  
Tanggal 24 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 : Nama Mahasiswa  
NIM : 19.11.3162

Menyatakan bahwa Skripsi dengan judul berikut:

**Analisis sentimen twitter tentang kenaikan bahan bakar minyak dengan metode naive bayes classifier (nbc)**

Dosen Pembimbing : Emi Seniwati, S.Kom, M.Cs

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, 24 Juni 2023

Yang Menyatakan,

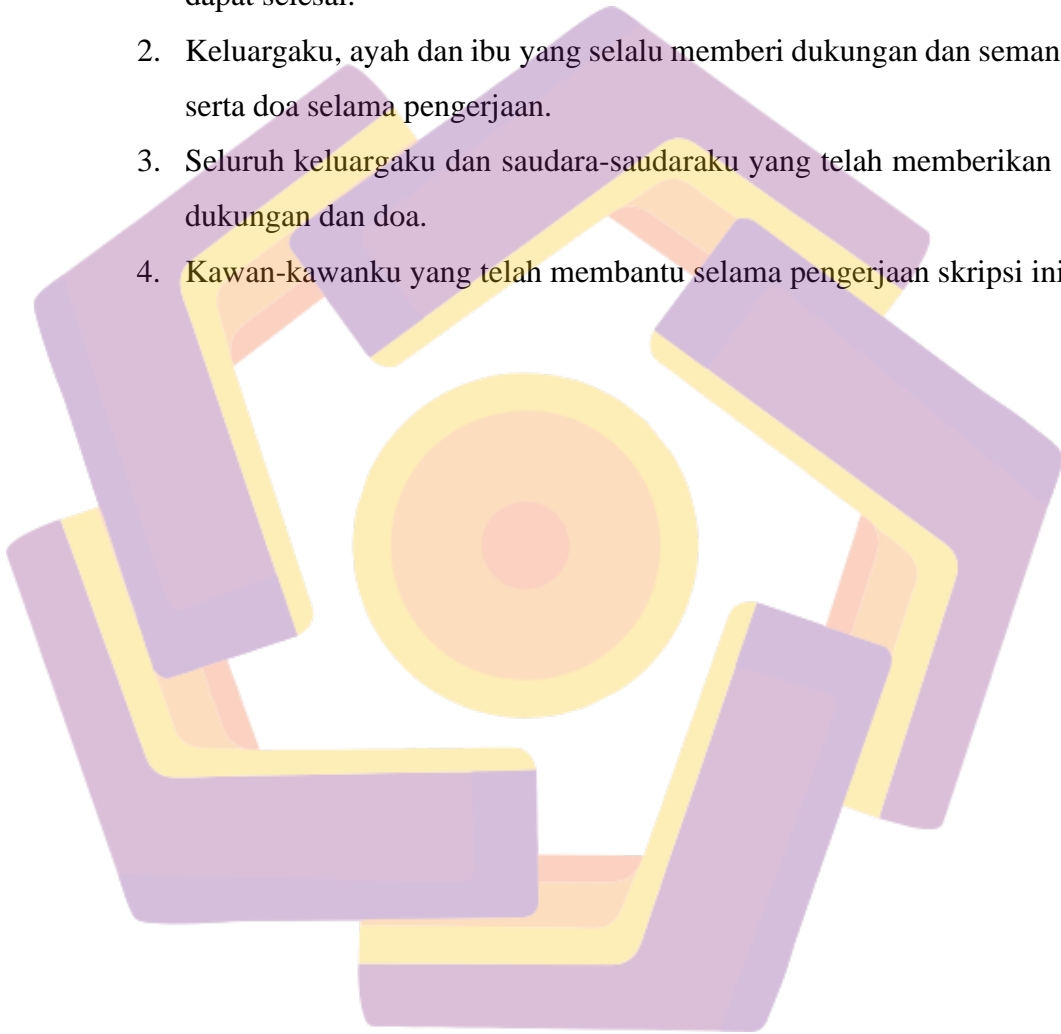


Mohammad Irham

## HALAMAN PERSEMBAHAN

Skripsi ini saya dipersembahkan kepada :

1. Allah SWT, karena hanya atas izin dan karunianya lah maka skripsi ini dapat selesai.
2. Keluargaku, ayah dan ibu yang selalu memberi dukungan dan semangat serta doa selama pengerjaan.
3. Seluruh keluargaku dan saudara-saudaraku yang telah memberikan dukungan dan doa.
4. Kawan-kawanku yang telah membantu selama pengerjaan skripsi ini.



## KATA PENGANTAR

Alhamdulillah segala puji bagi Allah Subhanahu wa Ta'ala yang telah melimpahkan rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul “Analisis Sentimen Twitter Tentang Kenaikan Bahan Bakar Minyak Menggunakan Naïve Bayes Clasifier (NBC)”. Skripsi ini ditujukan untuk memenuhi salah satu syarat kelulusan di program studi S1 Informatika Universitas AMIKOM Yogyakarta.

Penulis menyadari tanpa adanya dukungan beberapa pihak, penulis akan kesulitan dalam menyelesaikan skripsi ini. Maka dari itu, dengan penuh kerendahan hati, penulis ingin menyampaikan ucapan terima kasih kepada :

- Hanif Al Fatta, S. Kom., M. Kom. selaku Dekan Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta.
- Erni Seniwati, S. Kom., M. Cs dosen pembimbing yang telah membimbing penulis selama penyusunan skripsi ini.
- Dosen-dosen Universitas AMIKOM Yogyakarta yang telah menjadi penguji dalam sidang skripsi penulis.
- Dan semua pihak yang telah membantu penulis dalam menyusun skripsi ini yang tidak bisa penulis sebutkan satu persatu.

Akhir kata, penulis memahami betul bahwa tidak ada kesempurnaan melainkan milik Allah Subhanahu wa Ta'ala. Maka dari itu penulis meminta maaf atas kesalahan yang terdapat dalam skripsi ini. Meski begitu penulis tetap berharap bahwa skripsi ini dapat bermanfaat bagi pembaca pada umumnya serta dapat dijadikan sebagai referensi demi pengembangan ke arah yang lebih baik.

Yogyakarta, 24 juni 2023

Penulis

## 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.....	ix
DAFTAR GAMBAR.....	x
DAFTAR LAMPIRAN.....	xi
DAFTAR LAMBANG DAN SINGKATAN .....	xii
INTISARI .....	xiii
ABSTRACT.....	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah.....	2
1.3 Batasan Masalah .....	2
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 Analisis Sentimen .....	12
2.3 Naïve Bayes Clasifier .....	12
2.4 Preprocessing .....	13
2.5 Multinomial Naïve Bayes.....	14
2.6 Twitter .....	15
2.7 Python.....	15
BAB III METODE PENELITIAN .....	16



3.1	Gambaran Umum.....	16
3.2	Alur Penelitian .....	16
3.3	Analisis Kebutuhan.....	17
3.4	Analisi Data .....	17
3.4.1	Pengambilan Data .....	17
3.4.2	Cleaning data.....	18
3.4.3	Case Folding .....	19
3.4.4	Stopword Removal.....	20
3.4.5	Tokenizing .....	22
3.4.6	Stemming .....	23
3.4.7	Klasifikasi Naïve Bayes .....	25
3.4.8	Ekstraksi Fitur .....	25
3.4.9	Uji Model .....	25
3.4.10	Evaluasi.....	25
<b>BAB IV</b>	<b>HASIL DAN PEMBAHASAN .....</b>	<b>26</b>
4.1	Pengambilan Data.....	26
4.2	Preprocessing Data .....	27
4.3	Ekstraksi fitur .....	30
4.4	Klasifikasi Naïve Bayes .....	34
4.5	Uji Model.....	35
4.6	Evaluasi Model.....	36
<b>BAB V</b>	<b>PENUTUP .....</b>	<b>39</b>
5.1	Kesimpulan .....	39
5.2	Saran .....	39
	<b>DAFTAR PUSTAKA .....</b>	<b>40</b>
	<b>LAMPIRAN.....</b>	<b>42</b>
1.	Source Code .....	42

## DAFTAR TABEL

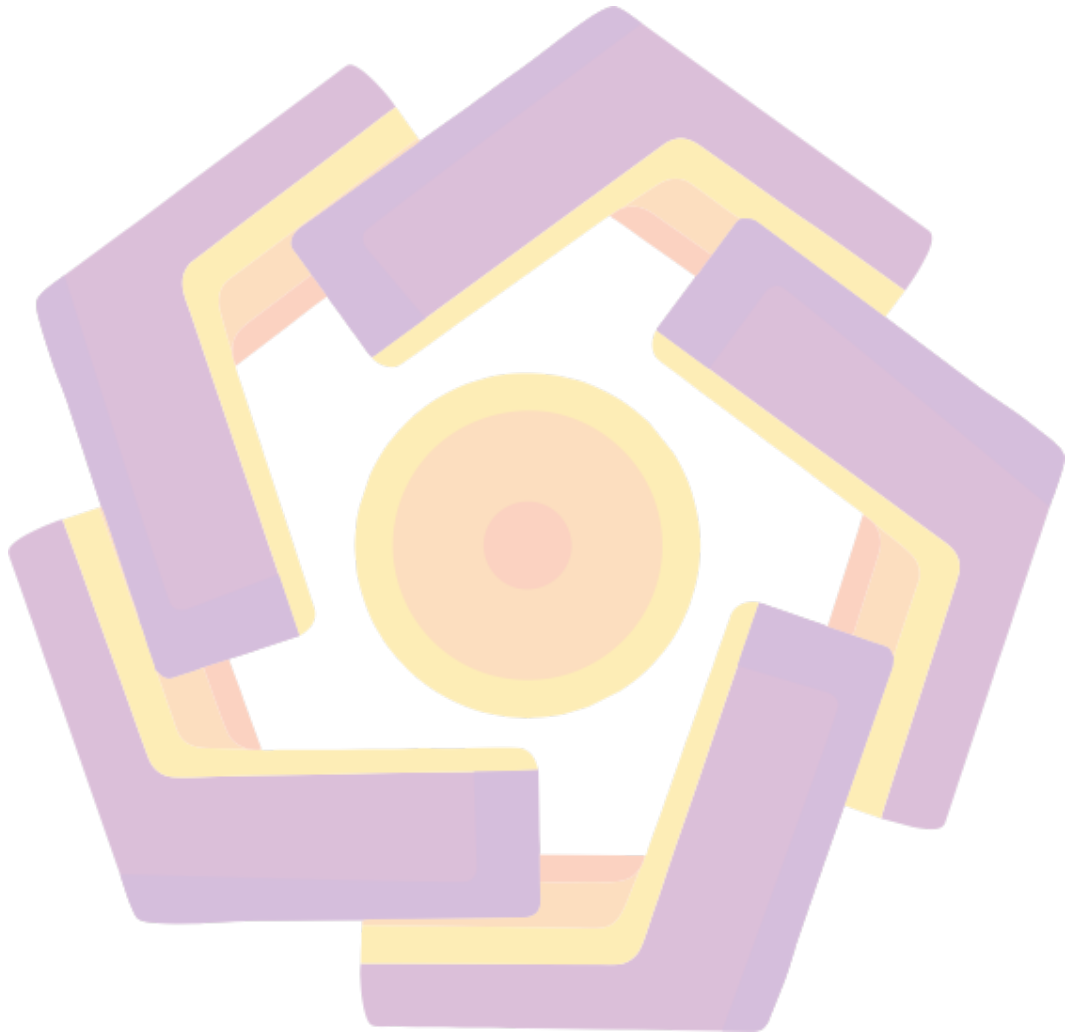
Tabel 2. 1 Keaslian Penelitian.....	7
Tabel 3. 1 Cleaning Data.....	18
Tabel 3. 2 Case Folding .....	19
Tabel 3. 3 Stopword Removal.....	21
Tabel 3. 4 Tokenisasi .....	22
Tabel 3. 5 Stemming Data.....	23
Tabel 4. 1 Pembuatan word vector.....	32
Tabel 4. 2 term frequency (TF).....	32
Tabel 4. 3 DF (Dokumen frequency).....	33
Tabel 4. 4 Perhitungan IDF.....	33
Tabel 4. 5 perhitungan TF-IDF .....	33
Tabel 4. 6 hasil pembobotan word vector .....	34
Tabel 4. 7 Model Confusion matrix .....	35
Tabel 4. 8 Hasil Confusion Matrix.....	37
Tabel 4. 9 Hasil Pengujian .....	37

## DAFTAR GAMBAR

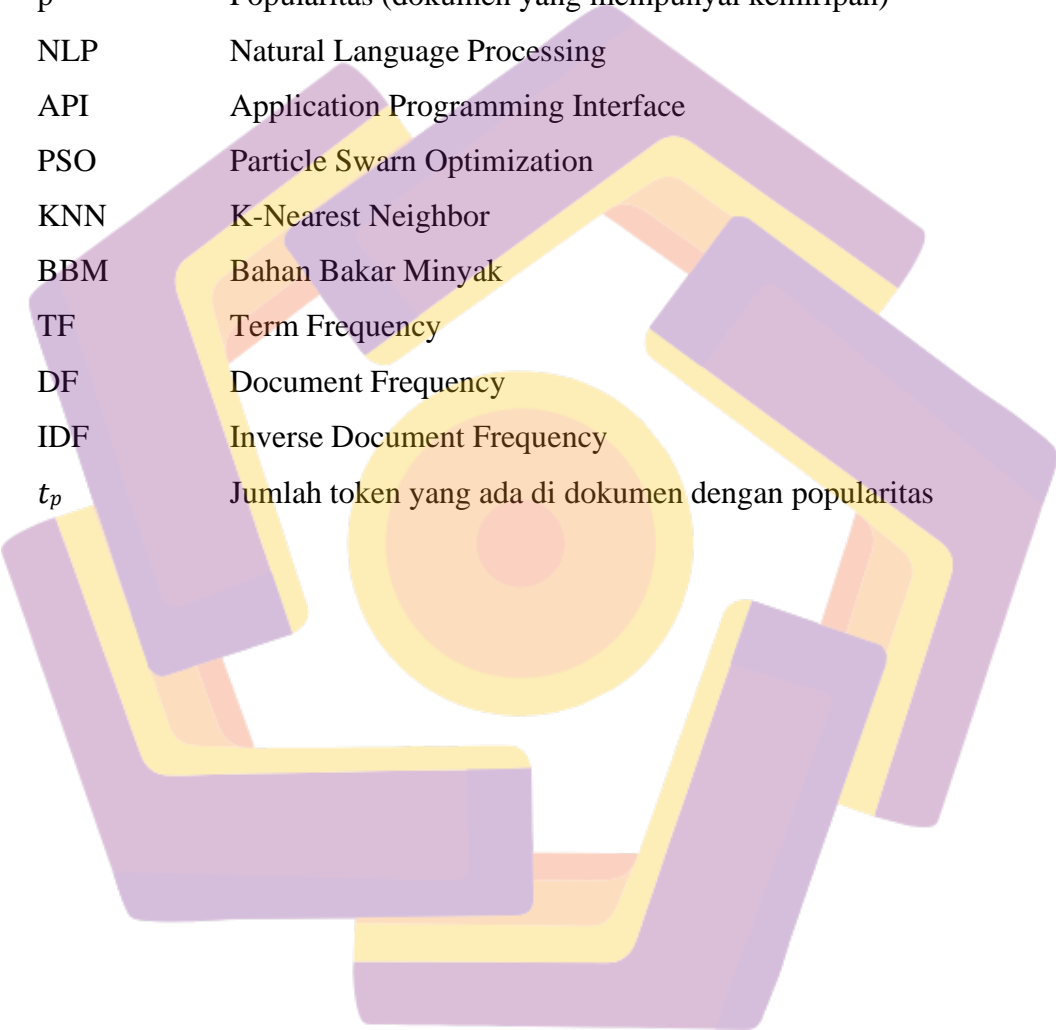
Gambar 3. 1 Alur Penelitian .....	16
Gambar 3. 2 hasil pengumpulan data.....	17
Gambar 4. 1 Proses Instalasi Snsrape.....	26
Gambar 4. 2 Proses deklarasi library .....	26
Gambar 4. 3 Kode program pengambilan data .....	27
Gambar 4. 4 hasil pengambilan data dan labeling data.....	27
Gambar 4. 5 kode program cleaning data .....	28
Gambar 4. 6 Deklarasi library stopwords.....	28
Gambar 4. 7 Implementasi kode program stoword removal.....	29
Gambar 4. 8 Implementasi kode program stopwords removal.....	29
Gambar 4. 9 implementasi proses stemming .....	30
Gambar 4. 10 hasil preprocessing data .....	30
Gambar 4. 11 hasil visualisasi kata.....	31
Gambar 4. 12 implementasi pembobotan kata.....	34
Gambar 4. 13 deklarasi library.....	34
Gambar 4. 14 implementasi naïve bayes .....	35
Gambar 4. 15 implementasi pembagian data.....	35
Gambar 4. 16 Perhitungan akurasi model.....	36
Gambar 4. 17 hasil uji model.....	36
Gambar 4. 18 Hasil pengujian.....	37

## DAFTAR LAMPIRAN

Source Code .....	39
-------------------	----



## DAFTAR LAMBANG DAN SINGKATAN



X	Data dokumen dengan kelas yang belum diketahui.
H	Hipotesis dari data X merupakan suatu kelas spesifik.
n	Jumlah dokumen
p	Popularitas (dokumen yang mempunyai kemiripan)
NLP	Natural Language Processing
API	Application Programming Interface
PSO	Particle Swarm Optimization
KNN	K-Nearest Neighbor
BBM	Bahan Bakar Minyak
TF	Term Frequency
DF	Document Frequency
IDF	Inverse Document Frequency
$t_p$	Jumlah token yang ada di dokumen dengan popularitas

## INTISARI

Bahan bakar minyak atau kita sebut BBM adalah komoditas penting bagi sebuah negara, bukan hanya untuk kendaraan transportasi saja tapi bisa menjadi barang dagangan untuk masyarakat. Bahan bakar minyak juga berperan penting dalam perkembangan dunia industri oleh sebab itu jika ada perubahan pada komoditas ini dapat menimbulkan perubahan dalam berbagai bidang. Oleh sebab itu pemerintah melakukan campur tangan dalam menentukan harga dan ketersediaan bahan bakar minyak. Bahan bakar minyak juga menjadi salah satu faktor yang dapat menyebabkan kenaikan bahan pokok dan inflasi pada sebuah negara, akibat hal itu banyak polemik yang muncul di masyarakat. Seiring perkembangan jaman dan teknologi, kini masyarakat dengan mudah menyampaikan opini mereka ke khalayak ramai tentang perubahan terjadi. Banyaknya opini yang beredar pada media social, tentang kenaikan bahan bakar minyak salah satunya pada media social twitter karena banyak opini pada media social tersebut, mengakibatkan sulitnya dalam indentifikasi apakah opini tersebut positif atau negatif. Maka dari itu dibuatlah penelitian ini. Penelitian ini bertujuan untuk mengidentifikasi apakah opini-opini yang beredar pada media social twitter dan juga menguji tingkat keakuratan pada sebuah metode algoritma. Algoritma yang digunakan pada penelitian ini adalah metode naïve bayes dengan hasil akurasi sebesar 81% dan dengan kesimpulan bahwa kenaikan bahan bakar minyak berdampak negatif pada masyarakat.

**Kata kunci:** Naïve Bayes, Analisa Sentimen, Confusion Matrix, Python.

## ABSTRACT

Fuel oil is an important commodity for a country, not only for transportation vehicles but can be merchandise for the community. Fuel oil also plays an important role in the development of world industry. Therefore, if there is a change in this commodity, it can cause changes in various fields. Therefore, the government intervenes in determining the price and availability of fuel oil. Fuel oil is also one of the factors that can cause an increase in staple goods and inflation in a country, due to the many polemics that have arisen in society. Along with the times and technology, now people can easily convey their opinions to the general public about the changes that have occurred. There are many opinions circulating on social media, regarding the increase in fuel oil, one of which is on social media Twitter because there are so many opinions on social media, making it difficult to identify whether these opinions are positive or negative. Therefore, this research was made. This study aims to identify whether there are opinions circulating on social media Twitter and also to test the level of accuracy of an algorithmic method. The algorithm used in this study is the naïve Bayes method with an accuracy of 81% and with the conclusion that the increase in fuel oil has a negative impact on society.

**Keyword:** Naïve Bayes, Sentiment Analysis, Confusion Matrix, Python.