

**IMPLEMENTASI ANALISIS SENTIMEN MASYARAKAT
MENGENAI KENAIKAN HARGA BBM DENGAN
METODE NAIVE BAYES**

SKRIPSI

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



disusun oleh

ARDIANSYAH

19.11.3026

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2023

**IMPLEMENTASI ANALISIS SENTIMEN MASYARAKAT
MENGENAI KENAIKAN HARGA BBM DENGAN METODE
NAIVE BAYES**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh
ARDIANSYAH
19.11.3026

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2023

HALAMAN PERSETUJUAN

SKRIPSI

**IMPLEMENTASI ANALISIS SENTIMEN MASYARAKAT MENGENAI
KENAIKAN HARGA BBM DENGAN METODE NAÏVE BAYES**

yang disusun dan diajukan oleh

Ardiansyah

19.11.3026

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 24 Agustus 2023

Dosen Pembimbing,



Nur'aini, M.Kom.

NIK. 190302066

HALAMAN PENGESAHAN

SKRIPSI

**IMPLEMENTASI ANALISIS SENTIMEN MASYARAKAT
MENGENAI KENAIKAN HARGA BBM DENGAN METODE
NAIVE BAYES**

yang disusun dan diajukan oleh

Ardiansyah

19.11.3026

Telah dipertahankan di depan Dewan Penguji
pada tanggal 24 Agustus 2023

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

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

Uyock Anggoro Saputro, S.Kom, M.Kom
NIK. 190302419

Nur'aini, M.Kom
NIK. 190302066



Skrripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 24 Agustus 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 : Ardiansyah
NIM : 19.11.3026

Menyatakan bahwa Skripsi dengan judul berikut:

Implementasi Analisis Sentimen Masyarakat Mengenai Kenaikan Harga BBM Dengan Metode Naïve Bayes

Dosen Pembimbing : Nur'aini 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, 18 Agustus 2023

Yang Menyatakan,



Ardiansyah

HALAMAN PERSEMBAHAN

Alhamdulillah, segala puji syukur penulis panjatkan kepada Allah SWT atas rahmat dan karunai serta hidayah-Nya, sehingga penulis dapat menyelesaikan tugas akhir skripsi dengan baik. Sholawat serta salam tidak lupa haturkan kepada Nabi Muhammad SAW yang telah membawa dunia ini ke jaman yang penuh rahmat serta penuh berkah. sehingga penulis mendapatkan kemudahan dan kelancaran dalam mengerjakan skripsi ini. Skripsi ini penulis persembahkan dengan rasa terima kasih kepada:

1. Orang tua serta keluarga yang selalu memberikan dukungan, motivasi dan doa restu tiada henti.
2. Ibu Nur'aini, M.Kom yang telah membimbing saya serta memberikan motivasi untuk selalu mengerjakan skripsi ini.
3. Teman-teman perjuangan 19 informatika 07 yang telah banyak memberikan semangat, berbagi keluh kesah selama melewati masa-masa perkuliahan yang penuh ada banyak permasalahan.
4. Sahabat-sahabat dari masa SMA yang selalu memberikan dukungan selama semasa menyelesaikan skripsi.
5. Teman-teman yang tidak bisa saya sebutkan namanya satu-persatu yang telah membantu dan men-support baik secara langsung maupun tidak langsung.

KATA PENGANTAR

Segala puji syukur kehadiran Allah swt. yang telah memberikan karunia, rahmat serta hidayah-nya sehingga penulis dapat menyelesaikan skripsi ini dengan baik dan maksimal. Skripsi yang berjudul **"Implementasi Analisis Sentimen Masyarakat Mengenal Kenalkan Harga Bbm Dengan Metode Naive Bayes"** dengan baik dan lancar. Laporan Skripsi ini disusun sebagai salah satu syarat dalam menyelesaikan masa studi program sarjana di Universitas AMIKOM Yogyakarta.

Dalam penulisan skripsi ini, penulis mengucapkan rasa hormat dan terima kasih atas motivasi, bimbingan, saran dan masukan dari berbagai pihak secara moral maupun spiritual. Pada kesempatan kali ini penulis mengucapkan terimakasih kepada:

1. Bapak Prof. Dr. M. Suyanto, M.M selaku Rektor Universitas AMIKOM Yogyakarta.
2. Bapak Hanif Al Fatta, M.Kom selaku Dekan Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.
3. Ibu Windha Mega Pradya D., M.Kom selaku Ketua Program Studi S1 Informatika Universitas Amikom Yogyakarta.
4. Ibu Nur'aini, M.Kom selaku Dosen Pembimbing yang selalu memberikan saran dan masukan dalam proses penulisan skripsi.

Penulis menyadari bahwa penelitian dan laporan Skripsi ini masih jauh dari sempurna. Oleh karena itu, kritik serta saran yang membangun sangat penulis harapkan demi kemajuan yang lebih baik di masa mendatang. semoga penelitian ini dapat memberikan tambahan ilmu pengetahuan dan manfaat pada semua pihak.

Yogyakarta, 24 Agustus 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	x
DAFTAR GAMBAR.....	xii
DAFTAR LAMBANG DAN SINGKATAN	xiii
DAFTAR ISTILAH	xiv
INTISARI	xv
ABSTRACT	xvi
BAB I PENDAHULUAN.....	1
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
1.6 Sistematika Penulisan	3
BAB II TINJAUAN PUSTAKA	5
2.1 Studi Literatur	5
2.2. Dasar Teori	12
2.2.1. Twitter	12
2.2.2. Analisis Sentimen	13
2.2.3. Application Programming Interface.....	13
2.2.4. Teks Mining	13
2.2.4.1. Clanning	14

2.2.4.2. Casefolding	14
2.2.4.3. Tokenize	14
2.2.4.4. Stopword Removal	14
2.2.4.5. Stemming	15
2.2.5. Term Frequency – Inverse Document Frequency	15
2.2.6. Naïve Bayes Classification	17
2.2.7. Confusion Matrix	18
BAB III METODE PENELITIAN	23
3.1. Alur Penelitian	23
3.2. Alat dan Bahan	24
3.2.1. Data Penelitian	24
3.2.2. Alat/instrumen	24
3.2.3. Kebutuhan Hardware / Perangkat Keras	24
3.2.4. Kebutuhan Software / Perangkat Lunak	24
3.3. Analisis Metode	25
3.3.1. Crawling Data	25
3.3.2. Preprocessing	26
3.3.2.1. Clanning	26
3.3.2.2. Casefolding	27
3.3.2.3. Tokenize	28
3.3.2.4. Stopword Removal	29
3.3.2.5. Stemming	30
3.3.3. Labelling Data	31
3.3.4. TF-IDF	32
3.3.5. Pembagian Data	32
3.3.6. Klasifikasi Naïve Bayes	33

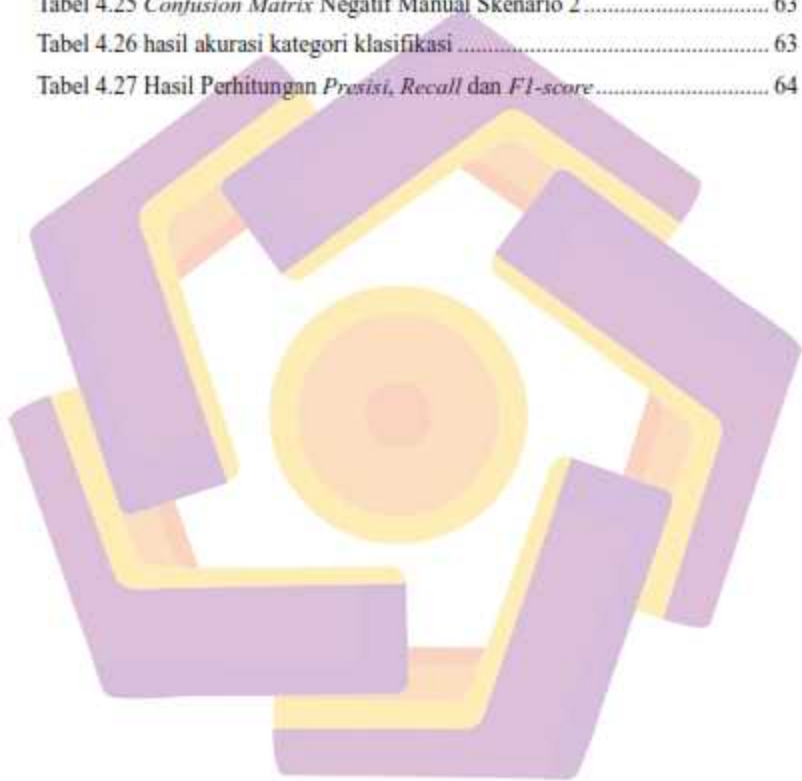
3.3.7. Evaluasi.....	33
BAB IV HASIL DAN PEMBAHASAN	34
4.1. Analisis Metode	34
4.1.1. Crawling Data	34
4.1.2. Preprocessing	35
4.1.2.1 Cleaning	35
4.1.2.2 Casefolding	36
4.1.2.3 Tokenize	37
4.1.2.4 Stopword Removal	37
4.1.2.5 Stemming	38
4.1.3. Labelling Data	39
4.1.4 TF-IDF	43
4.1.5 Pembagian Data	51
4.1.6 Klasifikasi Naïve Bayes	51
4.1.7 Evaluasi	54
BAB V PENUTUP.....	65
5.1 Kesimpulan	65
5.2 Saran	65
REFERENSI.....	67

s

DAFTAR TABEL

Tabel 2.1 Keaslian Penelitian.....	8
Tabel 2.2 Confusion Matrix.....	19
Tabel 3.1 Contoh <i>Crawling Data</i>	25
Tabel 3.2 Contoh <i>Cleaning Text</i>	26
Tabel 3.3 Contoh Casefolding.....	27
Tabel 3.4 Contoh <i>Tokenize</i>	28
Tabel 3.5 Contoh <i>Stopword Removal</i>	29
Tabel 3.6 Contoh <i>Stemming</i>	30
Tabel 3.7 Contoh Label Data.....	31
Tabel 3.8 Pembagian Skenario Data Latih dan Data Uji.....	33
Tabel 4.1 Contoh Klasifikasi Sentimen Positif.....	40
Tabel 4.2 Contoh Klasifikasi Sentimen Negatif.....	40
Tabel 4.3 Contoh Klasifikasi Sentimen Negatif.....	41
Tabel 4.4 Contoh Teks Sebelum Perhitungan Manual TF-IDF.....	43
Tabel 4.5 Tahapan Perhitungan Manual <i>Term Frequency</i>	44
Tabel 4.6 Perhitungan Manual <i>Inverse Document Frequency</i>	46
Tabel 4.7 Contoh Hasil Perhitungan Manual TF-IDF.....	48
Tabel 4.8 Pembagian Skenario.....	51
Tabel 4.9 Hasil Akurasi Berdasarkan Skenario.....	55
Tabel 4.10 Hasil Confusion Matrix Dataset Sistem Skenario 1.....	55
Tabel 4.11 Hasil Confusion Matrix Dataset Sistem Skenario 2.....	56
Tabel 4.12 Hasil Confusion Matrix Dataset Manual Skenario 1.....	56
Tabel 4.13 Hasil Confusion Matrix Dataset Manual Skenario 2.....	57
Tabel 4.14 <i>Confusion Matrix</i> Positif Prediksi Skenario 1.....	58
Tabel 4.15 <i>Confusion Matrix</i> Netral Prediksi Skenario 1.....	58
Tabel 4.16 <i>Confusion Matrix</i> Negatif Prediksi Skenario 1.....	59
Tabel 4.17 <i>Confusion Matrix</i> Positif Prediksi Skenario 2.....	59
Tabel 4.18 <i>Confusion Matrix</i> Netral Prediksi Skenario 2.....	60
Tabel 4.19 <i>Confusion Matrix</i> Negatif Prediksi Skenario 2.....	60

Tabel 4.20 <i>Confusion Matrix</i> Positif Manual Skenario 1	61
Tabel 4.21 <i>Confusion Matrix</i> Netral Manual Skenario 1	61
Tabel 4.22 <i>Confusion Matrix</i> Negatif Manual Skenario 1	61
Tabel 4.23 <i>Confusion Matrix</i> Positif Manual Skenario 2.....	62
Tabel 4.24 <i>Confusion Matrix</i> Netral Manual Skenario 2.....	62
Tabel 4.25 <i>Confusion Matrix</i> Negatif Manual Skenario 2.....	63
Tabel 4.26 hasil akurasi kategori klasifikasi	63
Tabel 4.27 Hasil Perhitungan <i>Precisi</i> , <i>Recall</i> dan <i>F1-score</i>	64

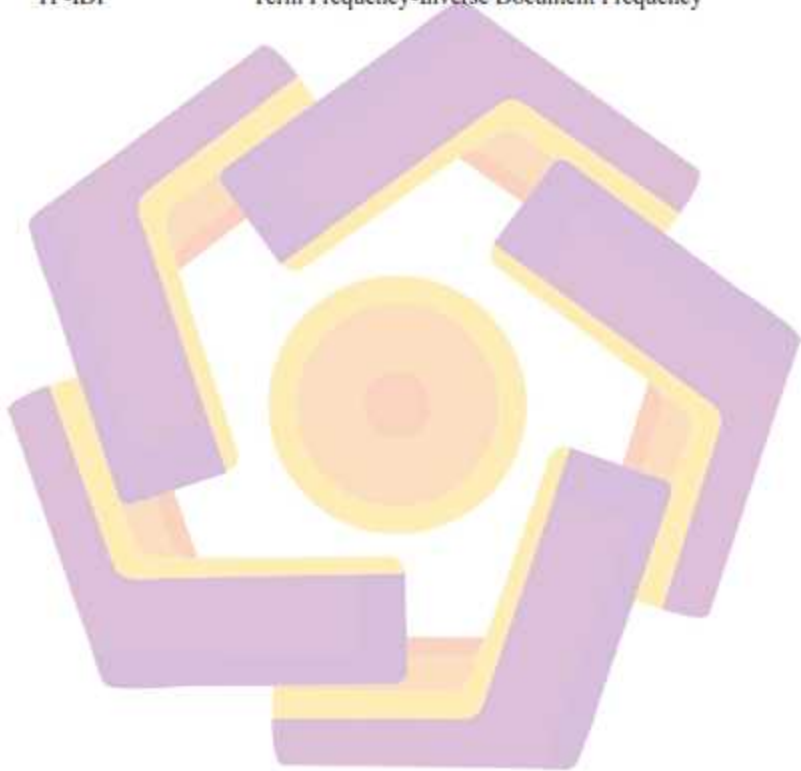


DAFTAR GAMBAR

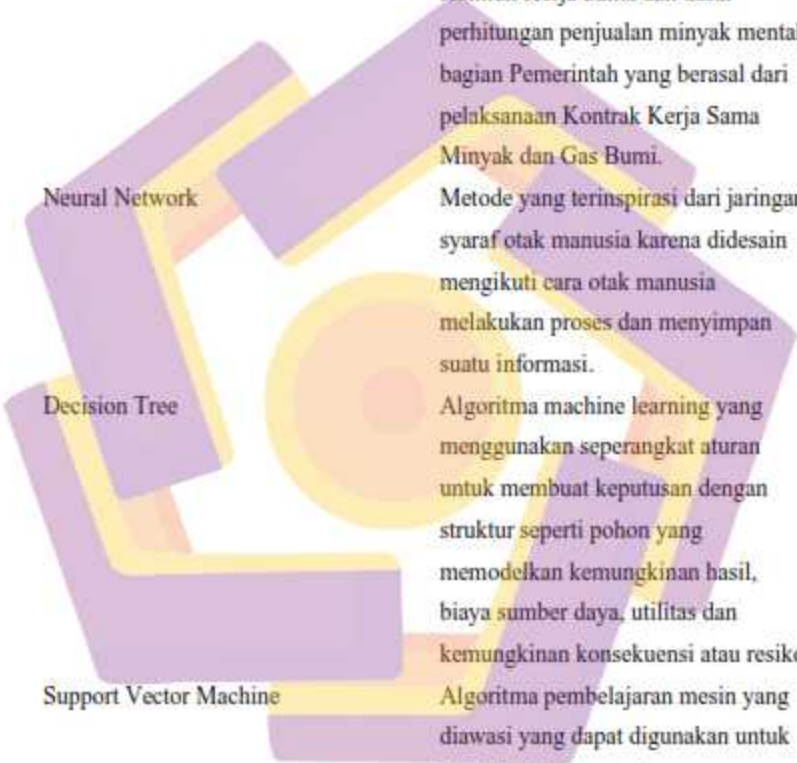
Gambar 3.1 Alur Penelitian.....	23
Gambar 4.1 Alur <i>Crawling</i> lewat RapidMiner	34
Gambar 4.2 Parameter pengaturan <i>Crawling</i> Twitter	34
Gambar 4.3 Pemanggilan <i>Library Cleaning</i>	35
Gambar 4.4 Proses <i>Cleaning</i> Teks	36
Gambar 4.5 Proses <i>Casefolding</i>	36
Gambar 4.6 Proses <i>Tokenize</i>	37
Gambar 4.7 Proses <i>Stopword Removal</i>	37
Gambar 4.8 Proses <i>Stopword Removal</i>	38
Gambar 4.9 Proses Pemanggilan <i>Library Sastrawi</i>	38
Gambar 4.10 Proses <i>Stemming</i>	39
Gambar 4.11 Diagram Hasil <i>Klasifikasi Labelling Data Manual</i>	43
Gambar 4.12 Pemanggilan library <i>TfidfVectorizer</i> dan <i>Spacy</i>	50
Gambar 4.13 Proses Eksekusi TF-IDF	50
Gambar 4.14 Hasil Pemrosesan TF-IDF.....	51
Gambar 4.15 Proses Pembagian Dataset.....	51
Gambar 4.16 Proses <i>Klasifikasi Naïve Bayes</i>	52
Gambar 4.17 Proses <i>Memastikan Data Bersih</i>	52
Gambar 4.18 Proses Pengaplikasikan <i>Sentimen Prediksi</i>	53
Gambar 4.19 Hasil Proses <i>Klasifikasi Sentimen</i>	53
Gambar 4.20 Diagram Jumlah <i>Sentimen Prediksi</i>	54
Gambar 4.21 Proses <i>Klasifikasi</i>	54
Gambar 4.22 Hasil Proses Alur <i>Klasifikasi</i>	55
Gambar 4.23 Proses <i>P्रेसisi, Recall, dan F1-Score</i>	64

DAFTAR LAMBANG DAN SINGKATAN

Vector	besaran yang mempunyai arah
NBC	Naïve Bayes Classifier
SVM	Support Vector Machine
TF-IDF	Term Frequency-Inverse Document Frequency



DAFTAR ISTILAH



ICP (Indonesia Crude Price)	Harga patokan minyak mentah Indonesia yang digunakan dalam penghitungan bagi hasil dalam Kontrak Kerja Sama dan dasar perhitungan penjualan minyak mentah bagian Pemerintah yang berasal dari pelaksanaan Kontrak Kerja Sama Minyak dan Gas Bumi.
Neural Network	Metode yang terinspirasi dari jaringan syaraf otak manusia karena didesain mengikuti cara otak manusia melakukan proses dan menyimpan suatu informasi.
Decision Tree	Algoritma machine learning yang menggunakan seperangkat aturan untuk membuat keputusan dengan struktur seperti pohon yang memodelkan kemungkinan hasil, biaya sumber daya, utilitas dan kemungkinan konsekuensi atau resiko.
Support Vector Machine	Algoritma pembelajaran mesin yang diawasi yang dapat digunakan untuk klasifikasi dan regresi.
Algoritma Nazief dan Adriani	Algoritma ini didasarkan pada aturan morfologi Bahasa Indonesia yang luas, yang dikumpulkan menjadi satu kelompok dan di enkapsulasi menjadi imbuhan yang diizinkan dan imbuhan yang tidak diizinkan.

INTISARI

Kenaikan harga BBM di tengah masyarakat banyak menuai pendapat pro dan kontra. Hal ini tidak terlepas dari mobilitas masyarakat yang begitu tinggi dalam penggunaan kendaraan pribadi serta berdampak pada sektor lain terutama kebutuhan pokok. Hal tersebut membuat *statement* masyarakat atas kebijakan pemerintah tersebut harus sesuai dengan keadaan di lapangan. Terlepas dari penyebab kenaikan harga BBM harus menyesuaikan harga minyak dunia. Pemerintah juga menanggarkan biaya untuk subsidi BBM namun, tidak selamanya menggunakan harga yang murah akan tetapi, pada waktu tertentu juga pemerintah akan menaikkan harga BBM agar tidak membenai anggaran APBN secara terus menerus. Menuai pendapat yang berisi negatif dan positif atas kebijakan tersebut. Berdasarkan permasalahan, penulis melakukan analisis sentimen tersebut menggunakan metode *Naïve Bayes*. Dengan menggunakan kumpulan dataset yang berisikan pendapat dari media *twitter*. Tahapan metode melalui dari *crawling* (pengumpulan data), *preprocessing* (pembersihan kata), TF-IDF (pembobotan kata), mengklasifikasikan dengan Algoritma *Naïve Bayes* dan terakhir evaluasi. Untuk melihat sentimen pendapat masyarakat kebijakan masyarakat yang menuai banyak tanggapan dan mengutarakan pendapat melalui *twitter*. Selain itu pada tahap evaluasi tersebut bertujuan seberapa besar nilai hubungan antar data uji dengan data latih serta tingkat efektif model klasifikasi *Naïve Bayes* terkait kenaikan harga BBM.

Kata Kunci: *Klasifikasi; Naïve Bayes; BBM.*

ABSTRACT

The increase in fuel prices in society reaps many opinions of the pros and cons. This is inseparable from the high mobility of people in the use of private vehicles and the impact on other sectors, especially basic needs. This makes the public's statement on the government's policy must be following the situation on the ground. Regardless of the cause of the increase in fuel prices must adjust to world oil prices. The government also budgets for fuel subsidies however, it doesn't always use cheap prices. However, at certain times the government will also increase fuel prices so as not to continuously burden the APBN budget. Reaping negative and positive opinions on this policy. Based on the problems, the authors conducted a sentiment analysis using the Naïve Bayes method. By using a collection of datasets that contain opinions from Twitter media. The stages of the method go through crawling (data collection), preprocessing (word cleaning), TF-IDF (word weighting), classifying with the Naïve Bayes Algorithm, and evaluation. To see the sentiments of public opinion, community policies that reap a lot of responses and express opinions through Twitter. In addition, the evaluation stage aims to determine the value of the relationship between the test data and the training data and the effectiveness of the Naïve Bayes classification model related to rising fuel prices.

Keywords: Classification; Naïve Bayes; BBM.