# ANALISIS SENTIMEN PUBLIK TERHADAP ISU KELANGKAAN BBM AKIBAT KONFLIK IRAN DAN AS DI MEDIA SOSIAL X MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER

**Proposal Tugas Akhir**

**Marcelino Felix Yussanto**

**NPM : 221711821**

**PROGRAM STUDI SISTEM INFORMASI**

**FAKULTAS TEKNOLOGI INDUSTRI**

**UNIVERSITAS ATMA JAYA YOGYAKARTA**

**2026**

# HALAMAN PENGESAHAN

Tugas Akhir Berjudul

ANALISIS SENTIMEN PUBLIK TERHADAP ISU KELANGKAAN BBM AKIBAT KONFLIK IRAN DAN AS DI MEDIA SOSIAL X MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER

yang disusun oleh

Marcelino Felix Yussanto

221711821

dinyatakan telah memenuhi syarat pada tanggal 24 Juli 2026

| | | Keterangan |
|---|---|---|
| Dosen Pembimbing 1 | : Elisabeth Marsella, S.S., M.Li. | Telah Menyetujui |
| Dosen Pembimbing 2 | : Rangga Perwiratama S.T., M.Kom. | Telah Menyetujui |
| Tim Penguji | | |
| Penguji 1 | : Elisabeth Marsella, S.S., M.Li. | Telah Menyetujui |
| Penguji 2 | : Dr. Fl. Sapty Rahayu, S.T., M.Kom. | Telah Menyetujui |
| Penguji 3 | : Julius Galih Prima Negara, S.Kom., S.A.P., M.Kom. | Telah Menyetujui |

Yogyakarta, 24 Juli 2026

Universitas Atma Jaya Yogyakarta

Teknologi Industri

Dekan

ttd.

Prof. Dr. Ir. Parama Kartika Dewa SP., S.T., M.T., IPU, ASEAN Eng.

Dokumen ini merupakan dokumen resmi UAJY yang tidak memerlukan tanda tangan karena dihasilkan secara elektronik oleh Sistem Bimbingan UAJY. UAJY bertanggung jawab penuh atas informasi yang tertera di dalam dokumen ini

## LEMBAR PERNYATAAN KEASLIAN PENELITIAN

Saya yang bertanda tangan di bawah ini:

| | |
|---|---|
| Nama Lengkap | : Marcelino Felix Yussanto |
| NPM | : 221711821 |
| Program Studi | : Sistem Informasi |
| Fakultas | : Teknologi Industri |
| Judul Penelitan | : Analisis Sentimen Publik Terhadap Isu Kelangkaan BBM Akibat Konflik Iran Dan As Di Media Sosial X Menggunakan Metode Naïve Bayes Classifier |

Menyatakan dengan ini:

1. Skripsi ini adalah benar merupakan hasil karya sendiri dan tidak merupakan salinan sebagian atau keseluruhan dari karya orang lain.
2. Memberikan kepada Universitas Atma Jaya Yogyakarta, berupa Hak Bebas Royalti non eksklusif (Non-Exclusive-Royalty-Free Right) atas Penelitian ini, dan berhak menyimpan, mengelola dalam pangkalan data, mendistribusikan, serta menampilkan untuk kepentingan akademis, tanpa perlu meminta izin selama tetap mencantumkan nama Peneliti.
3. Bersedia menanggung secara pribadi segala bentuk tuntutan hukum yang mengikuti atas pelanggaran Hak Cipta dalam pembuatan skripsi ini.

Demikianlah pernyataan ini dibuat dan dapat dipergunakan sebagaimana mestinya.

Yogyakarta, 21 Juli 2026

Yang menyatakan,



Marcelino Felix Yussanto

221711821

# PRAKATA

Puji dan syukur Peneliti panjatkan kepada Tuhan Yang Maha Esa atas berkat, rahmat, dan penyertaan-Nya, sehingga Peneliti dapat menyelesaikan tugas akhir yang berjudul "Analisis Sentimen Publik terhadap Isu Kelangkaan BBM Akibat Konflik Iran dan AS di Media Sosial X Menggunakan Metode Naïve Bayes Classifier" dengan baik. Tugas akhir ini disusun sebagai salah satu syarat untuk menyelesaikan studi pada Program Studi Sistem Informasi, Fakultas Teknologi Industri, Universitas Atma Jaya Yogyakarta.

Penyusunan tugas akhir ini tentu tidak lepas dari berbagai proses, tantangan, serta dukungan dari banyak pihak. Peneliti menyadari bahwa tanpa bantuan, arahan, dan doa dari berbagai pihak, penyusunan tugas akhir ini tidak akan dapat terselesaikan dengan baik. Oleh karena itu, Peneliti ingin menyampaikan rasa terima kasih yang sebesar-besarnya kepada:

1. Orang tua dan keluarga Peneliti yang selalu memberikan doa, dukungan, semangat, serta motivasi selama proses penyusunan tugas akhir ini.
2. Ibu Elisabeth Marsella, S.S., M.Li., selaku Dosen Pembimbing 1 yang telah memberikan arahan, bimbingan, masukan, serta waktu selama proses penyusunan tugas akhir ini.
3. Bapak Rangga Perwiratama, S.T., M.Kom., selaku Dosen Pembimbing 2 yang telah membantu memberikan bimbingan, saran, dan masukan yang sangat berarti dalam proses penelitian ini.
4. Pasangan Peneliti yang selalu memberikan dukungan, semangat, pengertian, serta motivasi selama Peneliti menjalani proses penyusunan tugas akhir ini.
5. Teman-teman seperjuangan dan rekan satu angkatan yang telah memberikan dukungan, serta semangat kebersamaan selama masa perkuliahan hingga proses penyusunan tugas akhir ini.
6. Seluruh pihak yang tidak dapat disebutkan satu per satu, yang telah memberikan bantuan, doa, dan dukungan kepada Peneliti dalam menyelesaikan tugas akhir ini.

# INTISARI

Penelitian ini bertujuan untuk menganalisis sentimen publik terhadap isu kelangkaan bahan bakar minyak (BBM) akibat konflik Iran dan Amerika Serikat di media sosial X menggunakan metode Naïve Bayes Classifier. Data penelitian diperoleh melalui proses *scraping* pada media sosial X dengan periode pengambilan data 1 Maret 2026 hingga 31 Maret 2026. Data awal yang berhasil dikumpulkan berjumlah 4.743 *tweet*, kemudian melalui proses seleksi dan *preprocessing* sehingga diperoleh 3.467 *tweet* yang digunakan dalam penelitian. Selanjutnya, data diberi label sentimen positif, negatif, dan netral, kemudian dilakukan pembobotan kata menggunakan metode TF-IDF sebelum diklasifikasikan menggunakan algoritma Multinomial Naïve Bayes. Hasil penelitian menunjukkan bahwa sentimen negatif lebih dominan dibandingkan sentimen positif, dan sentimen netral, dengan jumlah 2.947 *tweet* termasuk dalam sentimen negatif, 399 *tweet* termasuk dalam sentimen positif, dan 121 *tweet* termasuk dalam sentimen netral. Model Naïve Bayes Classifier memperoleh akurasi sebesar 86%, dengan performa yang lebih baik dalam mengenali sentimen negatif dibandingkan sentimen positif, dan sentimen netral. Dominasi sentimen negatif menunjukkan bahwa masyarakat cenderung memiliki kekhawatiran terhadap isu kelangkaan BBM, terutama berkaitan dengan kenaikan harga, ketersediaan stok, subsidi, dan dampak ekonomi. Penelitian ini diharapkan dapat memberikan gambaran mengenai persepsi publik di media sosial serta menjadi referensi dalam pengembangan penelitian analisis sentimen dan text mining.

**Kata Kunci**: Analisis Sentimen, Kelangkaan BBM, Media Sosial X, TF-IDF, Naïve Bayes Classifier.

# *ABSTRACT*

*This study aims to analyze public sentiment toward the issue of fuel oil scarcity caused by the conflict between Iran and the United States on social media X using the Naïve Bayes Classifier method. The research data were obtained through a scraping process on social media X, with the data collection period from March 1, 2026, to March 31, 2026. The initial data collected consisted of 4,743 tweets, which were then selected and preprocessed, resulting in 3.467 tweets used in this study. Furthermore, the data were labeled into positive, negative, and neutral sentiments, followed by word weighting using the TF-IDF method before being classified using the Multinomial Naïve Bayes algorithm. The results showed that negative sentiment was more dominant than positive and neutral sentiments, with 2.947 tweets classified as negative sentiment, 399 tweets as positive sentiment, and 121 tweets as neutral sentiment. The Naïve Bayes Classifier model achieved an accuracy of 86%, with better performance in identifying negative sentiment compared to positive and neutral sentiments. The dominance of negative sentiment indicates that the public tends to have concerns regarding the issue of fuel scarcity, particularly related to price increases, stock availability, subsidies, and economic impacts. This study is expected to provide an overview of public perception on social media and serve as a reference for the development of sentiment analysis and text mining research.*



***Keywords:*** *Sentiment Analysis, Fuel Scarcity, Social Media X, TF-IDF, Naïve Bayes Classifier.*

# DAFTAR ISI

HALAMAN PENGESAHAN ..... i
LEMBAR PERNYATAAN KEASLIAN PENELITIAN ..... ii
PRAKATA ..... iii
INTISARI ..... iv
ABSTRACT ..... v
DAFTAR ISI ..... vi
DAFTAR GAMBAR ..... viii
DAFTAR TABEL ..... ix

BAB I PENDAHULUAN ..... 1
1.1 Latar Belakang ..... 1
1.2 Rumusan Masalah ..... 4
1.3 Pertanyaan Penelitian ..... 4
1.4 Batasan Masalah ..... 5
1.5 Tujuan Penelitian ..... 5
1.6 Manfaat Penelitian ..... 5
1.7 Bagan Keterkaitan ..... 6
BAB II TINJAUAN PUSTAKA ..... 7
2.1 Studi Sebelumnya ..... 7
2.2 Dasar Teori ..... 13
2.1.1 Analisis Sentimen ..... 13
2.1.2 Text Mining ..... 13
2.1.3 Media Sosial ..... 14
2.1.4 X ..... 14
2.1.5 TF-IDF ..... 15
2.1.6 Naïve Bayes Classifier (NBC) ..... 16
2.1.7 Confusion Matrix ..... 18
BAB III METODOLOGI PENELITIAN ..... 19
3.1 Tahapan Penelitian ..... 19
3.1.1 Identifikasi Masalah ..... 19
3.1.2 Studi Literatur ..... 20
3.1.3 Pengumpulan Data ..... 20
3.1.4 Preprocessing Data ..... 21
3.1.5 Pelabelan Sentimen ..... 23
3.1.6 Ekstraksi Fitur Menggunakan TF-IDF ..... 24
3.1.7 Klasifikasi Menggunakan Naïve Bayes ..... 24

3.1.8 Evaluasi Model ..... 24
3.1.9 Interpretasi Hasil ..... 25
3.1.10 Visualisasi Data ..... 25
3.2 Perangkat dan Tools Penelitian ..... 26
BAB IV HASIL DAN PEMBAHASAN ..... 26
4.1 Pengumpulan Data ..... 26
4.2 *Preprocessing* Data ..... 29
4.2.1 *Cleaning* ..... 29
4.2.2 Case Folding ..... 35
4.2.3 Normalization ..... 36
4.2.4 *Tokenizing* ..... 38
4.2.5 *Stopword removal* ..... 40
4.2.6 *Stemming* ..... 42
4.3 Pelabelan Sentimen ..... 44
4.4 Pembobotan TF-IDF ..... 47
4.5 Klasifikasi ..... 50
4.5.1 Naïve Bayes Classifer (NBC) ..... 50
4.6 Evaluasi Model ..... 52
4.7 Interpretasi Hasil ..... 58
4.8 Visualisasi Data ..... 61
BAB V KESIMPULAN DAN SARAN ..... 67
5.1 Kesimpulan ..... 67
5.2 Saran ..... 67
DAFTAR PUSTAKA ..... 69
LAMPIRAN ..... 74

# DAFTAR GAMBAR

Gambar 1.1 Bagan Keterkaitan ..... 6
Gambar 3.1 Flowchart Tahapan Penelitian ..... 19
Gambar 4.1 Grafik Presentase Hasil Sentimen ..... 47
Gambar 4.2 Top 10 Kata dengan TF-IDF Tertinggi ..... 49
Gambar 4.3 Confusion Matrix Naïve Bayes ..... 53
Gambar 4.4 Confusion Matrix Setelah Menggunakan SMOTE ..... 54
Gambar 4.5 Learning curve Naïve Bayes ..... 56
Gambar 4.6 Word Cloud Sentimen ..... 61
Gambar 4.7 Frekuensi kata ..... 62
Gambar 4.8 World Cloud Sentimen Positif ..... 63
Gambar 4.9 Word Cloud Sentimen Negatif ..... 64
Gambar 4.10 Tabel WordCloud Sentimen Netral ..... 65

# DAFTAR TABEL

Tabel 2.1 Studi Sebelumnya ..... 12
Tabel 2.2 Tabel Perbandingan Kelebihan dan Kekurangan Beberapa Algoritma ..... 17
Tabel 4.1 Hasil Pengumpulan Data ..... 28
Tabel 4.2 Tabel Kode Proses Web Scraping ..... 29
Tabel 4.3 Tabel Kode Filtering Data Komentar Berbahasa Indonesia ..... 30
Tabel 4.4 Tabel kode Hapus Data Duplikat ..... 31
Tabel 4.5 Tabel Contoh Tweet Tidak Relevan pada Tahap Cleaning Manual ..... 33
Tabel 4.6 Tabel Kode Hapus Karakter Khusus pada kolom Comment ..... 34
Tabel 4.7 Hasil Cleaning ..... 35
Tabel 4.8 Tabel Kode Case Folding ..... 36
Tabel 4.9 Hasil Cleaning ..... 36
Tabel 4.10 Tabel Kode Normalization ..... 38
Tabel 4.11 Hasil Normalization ..... 38
Tabel 4.12 Tabel Kode Tokenizing ..... 39
Tabel 4.13 Hasil Tokenizing ..... 39
Tabel 4.14 Tabel Kode Stopword Removal ..... 41
Tabel 4.15 Hasil Stopword removal ..... 42
Tabel 4.16 Tabel Kode Stemming ..... 43
Tabel 4.17 Hasil Stemming ..... 44
Tabel 4.18 Tabel Kode Pelabelan Sentimen ..... 46
Tabel 4.19 Data Setiap Kelas Sentimen ..... 47
Tabel 4.20 Tabel Kode Pembobotan dengan TF-IDF ..... 49
Tabel 4.21 Tabel Kode Klasifikasi Naïve Bayes ..... 51
Tabel 4.22 Hasil Perhitungan Naïve Bayes Classifer ..... 51
Tabel 4.23 Tabel Hasil Perhitungan Menggunakan SMOTE ..... 54
Tabel 4.24 Tabel Perbandingan Akurasi dan Waktu Model beberapa Algoritma ..... 57
Tabel 4.25 Tabel Perbandingan Naïve Bayes Sebelum dan Sesudah SMOTE ..... 59