

ABSTRAK

Nama : Ahmad Aditia Subekti

Npm : 2255201110

Pembimbing : Muhammad Husni Rifqo, M.Kom.

Kemajuan teknologi informasi dan media sosial telah menjadikan platform X sebagai salah satu saluran yang paling banyak digunakan oleh masyarakat untuk menyampaikan pandangan mereka mengenai berbagai kebijakan pemerintah, termasuk kebijakan terkait kenaikan harga BBM. Banyaknya opini yang dipublikasikan setiap hari menghasilkan kumpulan data yang sangat besar, sehingga diperlukan teknik analisis otomatis untuk mengidentifikasi tren sentimen publik secara efektif. Penelitian ini bertujuan untuk mengkaji sentimen publik terhadap kebijakan kenaikan harga BBM yang diumumkan oleh Menteri Energi dan Sumber Daya Mineral (ESDM), Bahlil Lahadalia, dengan menerapkan algoritma Logistic Regression. Penelitian ini dilakukan berdasarkan kerangka kerja Cross-Industry Standard Process for Data Mining (CRISP-DM). Data dikumpulkan menggunakan teknik web scraping pada platform X melalui layanan Apify, yang diintegrasikan dengan bahasa pemrograman Python. Proses pengumpulan data tersebut menghasilkan total 2.262 cuitan. Selanjutnya, data tersebut melalui beberapa langkah prapemrosesan yaitu, pembersihan, tokenisasi, penghapusan stopword, normalisasi, dan stemming sehingga menghasilkan 1.791 tweet yang layak untuk dianalisis. Data yang telah diproses kemudian dikategorikan ke dalam tiga kelas sentimen: positif, negatif, dan netral. Langkah selanjutnya adalah ekstraksi fitur menggunakan metode *Term Frequency Inverse Document Frequency* (TF-IDF); dataset tersebut kemudian dibagi menjadi set pelatihan dan set uji dengan perbandingan 80:20. Model Logistic Regression kemudian dievaluasi menggunakan Confusion Matrix, dengan akurasi, precision, recall, dan F1-score sebagai metrik kinerja. Berdasarkan hasil pengujian, 195 titik data diidentifikasi sebagai sentimen positif, 1.358 sebagai sentimen negatif, dan 238 sebagai sentimen netral. Model tersebut mencapai tingkat akurasi sebesar 77,72%, yang menunjukkan bahwa algoritma Logistic Regression cukup efektif dalam mengklasifikasikan sentimen publik. Dominasi sentimen negatif menunjukkan bahwa mayoritas pengguna platform X mengungkapkan respons yang sebagian besar tidak mendukung terhadap kebijakan kenaikan harga bahan bakar. Hasil penelitian ini diharapkan dapat memberikan wawasan bagi pemerintah dalam memahami persepsi publik terhadap kebijakan yang diterapkan serta menjadi acuan bagi penelitian selanjutnya terkait analisis sentimen berbasis pembelajaran mesin.

Kata kunci: Analisis Sentimen, *Logistic Regression*, Media Sosial X, TF-IDF, Kenaikan Harga BBM.

ABSTRACT

SENTIMENT ANALYSIS OF SOCIAL MEDIA COMMENTS ON X REGARDING THE FUEL PRICE INCREASE POLICY USING THE LOGISTIC REGRESSION ALGORITHM

NAME : Ahmad Aditia Subekti
STUDENT ID : 2255201110
SUPERVISOR : Muhammad Husni Rifqo, M.Kom.

The advancement of information technology and social media has made the X platform one of the most widely used channels for people to express their opinions on various government policies, including policies related to fuel price increases. The large volume of opinions published daily generates a substantial amount of data, requiring automated analysis techniques to effectively identify trends in public sentiment. This study aims to examine public sentiment toward the fuel price increase policy announced by the Minister of Energy and Mineral Resources (ESDM), Bahlil Lahadalia, by applying the Logistic Regression algorithm. This research was conducted based on the Cross-Industry Standard Process for Data Mining (CRISP-DM) framework. Data were collected through web scraping on the X platform using the Apify service, integrated with the Python programming language. The data collection process resulted in a total of 2,262 tweets. The data then underwent several preprocessing stages, including cleaning, tokenization, stopword removal, normalization, and stemming, resulting in 1,791 tweets suitable for analysis. The processed data were categorized into three sentiment classes: positive, negative, and neutral. The next stage involved feature extraction using the Term Frequency-Inverse Document Frequency (TF-IDF) method. The dataset was then divided into training and testing sets using an 80:20 ratio. The Logistic Regression model was evaluated using a Confusion Matrix, with accuracy, precision, recall, and F1-score as performance metrics. Based on the test results, 195 data points were identified as positive sentiment, 1,358 as negative sentiment, and 238 as neutral sentiment. The model achieved an accuracy of 77.72%, indicating that the Logistic Regression algorithm was sufficiently effective in classifying public sentiment. The dominance of negative sentiment indicates that the majority of users on the X platform expressed predominantly unfavorable responses to the fuel price increase policy. The results of this study are expected to provide insights for the government in understanding public perceptions of implemented policies and serve as a reference for future research on machine learning-based sentiment analysis.

Keywords: Sentiment Analysis, Logistic Regression, X Social Media, TF-IDF, Fuel Price Increase

BAB I

PENDAHULUAN

1.1 Latar Belakang

Era digital ditandai dengan kemajuan teknologi informasi dan komunikasi yang mengubah secara mendasar cara masyarakat mencari, mengakses, hingga menyebarluaskan informasi. Salah satu bentuk perubahan tersebut terlihat dari semakin banyaknya masyarakat yang menggunakan media sosial sebagai wadah komunikasi untuk mengutarakan pendapat, kritik, atau tanggapan atas berbagai persoalan, baik yang menyangkut isu sosial, politik, maupun kebijakan pemerintah. Salah satu platform media sosial yang cukup banyak digunakan adalah X (dulu bernama Twitter), sebab platform ini dapat menyebarkan informasi dengan cepat dan bersifat real-time. Besarnya jumlah pengguna X di Indonesia membuat platform ini menjadi sumber data yang cukup relevan untuk mempelajari opini masyarakat terhadap berbagai kebijakan pemerintah. Unggahan maupun komentar dari pengguna dapat dijadikan bahan untuk mengetahui kecenderungan opini masyarakat secara lebih cepat dan menyeluruh terhadap isu yang sedang hangat dibicarakan (Kurniawan Santoso, 2022a). Kebijakan penyesuaian harga Bahan Bakar Minyak (BBM) termasuk salah satu kebijakan pemerintah yang berdampak besar bagi perekonomian dan kehidupan masyarakat. Naik turunnya harga BBM bukan hanya memengaruhi biaya transportasi, melainkan juga dapat mendorong naiknya harga kebutuhan pokok lain serta mengganggu kestabilan ekonomi secara luas. Karena itu, setiap keputusan terkait BBM hampir selalu menarik perhatian publik dan memicu berbagai macam reaksi. Salah satu kebijakan yang cukup ramai dibicarakan yaitu penyesuaian harga BBM yang diumumkan oleh Menteri Energi

dan Sumber Daya Mineral (ESDM), Bahlil Lahadalia. Pengumuman tersebut menuai berbagai reaksi masyarakat, mulai dari yang mendukung, mengkritik, sampai menolak, yang sebagian besar disuarakan lewat media sosial. Kondisi ini memperlihatkan bahwa media sosial kini menjadi salah satu ruang utama bagi masyarakat untuk mengungkapkan opini, harapan, dan tanggapan atas kebijakan pemerintah (Khatib Sulaiman et al., 2023a).

Banyaknya komentar yang muncul dari pengguna media sosial membuat analisis opini secara manual menjadi kurang praktis, sebab membutuhkan waktu, tenaga, dan sumber daya yang tidak sedikit. Untuk mengatasinya, dibutuhkan metode otomatis yang dapat mengelompokkan opini masyarakat ke dalam kategori sentimen positif, negatif, atau netral. Salah satu metode yang sering digunakan untuk keperluan tersebut adalah analisis sentimen, yaitu teknik dalam bidang text mining dan Natural Language Processing (NLP) yang berguna untuk mengenali dan mengklasifikasikan kecenderungan opini dari data berbentuk teks. Penggunaan analisis sentimen pada data media sosial memungkinkan proses pengenalan persepsi masyarakat terhadap suatu kebijakan berlangsung lebih cepat, sistematis, dan terstruktur.

Tingkat keberhasilan analisis sentimen sangat bergantung pada metode klasifikasi yang dipakai untuk mengolah data teks. Salah satu algoritma machine learning yang cukup sering dipakai untuk klasifikasi adalah Logistic Regression, yang bekerja dengan cara menghitung nilai probabilitas suatu data sebelum menetapkan kelas yang paling tepat. Dalam berbagai penelitian analisis sentimen, Logistic Regression dikenal memiliki beberapa kelebihan, di antaranya proses komputasi yang cukup cepat, cara implementasi yang tidak rumit, serta kemampuan

menghasilkan klasifikasi yang cukup baik. Algoritma ini juga cocok digunakan untuk mengolah data berdimensi tinggi, seperti kumpulan komentar atau cuitan di platform X, sehingga sering dipakai dalam berbagai penelitian analisis sentimen (Lidinillah et al., 2023).

Beberapa penelitian sebelumnya membuktikan bahwa algoritma Logistic Regression cukup handal digunakan untuk analisis sentimen pada data media sosial. Penelitian Kurniawan Santoso (2022a), misalnya, mendapati bahwa algoritma ini menghasilkan tingkat klasifikasi yang lebih baik pada data Twitter berbahasa Indonesia jika dibandingkan dengan metode klasifikasi lain. Temuan yang tidak jauh berbeda juga diungkapkan oleh Lidinillah et al. (2023), yang menyebutkan bahwa Logistic Regression memiliki kinerja yang cukup bersaing dalam mengelompokkan sentimen pada data media sosial. Mengacu pada berbagai hasil penelitian tersebut, dapat dikatakan bahwa Logistic Regression merupakan salah satu algoritma yang cocok dan layak dipakai untuk penelitian analisis sentimen berbasis media sosial. Atas dasar itu, penelitian ini memakai algoritma Logistic Regression untuk menganalisis sentimen masyarakat terhadap kebijakan kenaikan harga BBM yang diumumkan Menteri ESDM, dengan data yang diambil dari platform X. Melalui metode tersebut, diharapkan diperoleh gambaran yang akurat mengenai kecenderungan opini publik atas kebijakan tersebut.

Walaupun penelitian tentang analisis sentimen sudah cukup banyak dilakukan, masih ada beberapa sisi yang belum banyak diulas. Kebanyakan penelitian sebelumnya lebih menyoroti topik hiburan, aplikasi digital, atau isu sosial secara umum. Sementara itu, kajian mengenai sentimen masyarakat terhadap kebijakan kenaikan harga Bahan Bakar Minyak (BBM) yang diumumkan Menteri Energi dan

Sumber Daya Mineral (ESDM), Bahlil Lahadalia, masih terbilang jarang. Selain itu, penelitian-penelitian sebelumnya cenderung lebih menitikberatkan pada perbandingan performa berbagai algoritma klasifikasi, tanpa menelaah lebih dalam pola dan dinamika opini publik terhadap kebijakan sektor energi. Padahal, kebijakan penyesuaian harga BBM adalah isu strategis yang berdampak luas pada perekonomian dan kehidupan masyarakat Indonesia. Oleh sebab itu, dibutuhkan penelitian yang secara khusus mengkaji sentimen masyarakat terhadap kebijakan tersebut menggunakan data dari platform media sosial X, agar diperoleh gambaran yang lebih aktual dan lengkap mengenai tanggapan publik terhadap kebijakan pemerintah (Khatib Sulaiman et al., 2023b).

Bertolak dari persoalan yang telah dipaparkan, penelitian ini dimaksudkan untuk menganalisis sentimen masyarakat terhadap kebijakan kenaikan harga Bahan Bakar Minyak (BBM) yang diumumkan oleh Menteri Energi dan Sumber Daya Mineral (ESDM), Bahlil Lahadalia, dengan memanfaatkan komentar-komentar yang diunggah di platform media sosial X serta menerapkan algoritma Logistic Regression. Penelitian ini diharapkan bisa memberikan gambaran tentang kecenderungan opini publik terhadap kebijakan tersebut, baik yang bersifat positif, negatif, maupun netral. Selain itu, hasil penelitian ini diharapkan mampu menambah wawasan mengenai penerapan metode machine learning, terutama algoritma Logistic Regression, dalam bidang analisis sentimen di Indonesia. Secara praktis, temuan penelitian ini diharapkan bisa menjadi bahan pertimbangan bagi pemerintah untuk memahami tanggapan masyarakat terkait kebijakan di sektor energi, sehingga dapat mendukung proses evaluasi dan penyusunan kebijakan publik yang lebih tepat sasaran.

1.2 Rumusan Masalah

Mengacu pada latar belakang yang sudah dipaparkan, rumusan masalah penelitian ini yaitu:

- a. Bagaimana kecenderungan sentimen masyarakat di media sosial X terhadap kebijakan kenaikan harga BBM?
- b. Bagaimana kemampuan algoritma Logistic Regression dalam mengklasifikasikan sentimen komentar masyarakat di media sosial X terkait kebijakan kenaikan harga BBM menjadi kategori positif, negatif, dan netral?
- c. Bagaimana kinerja algoritma Logistic Regression dalam melakukan klasifikasi sentimen opini masyarakat terhadap perubahan harga BBM berdasarkan nilai accuracy, precision, recall, F1-score, dan confusion matrix?

1.3 Tujuan Penelitian

Adapun tujuan dari penelitian ini yaitu:

- a. Menganalisis sentimen komentar pada platform media sosial X mengenai kebijakan kenaikan harga bahan bakar yang disampaikan oleh Bahlil Lahadalia dengan menggunakan algoritma Logistic Regression.
- b. Menjelaskan kecenderungan opini publik terhadap kebijakan kenaikan harga bahan bakar sekaligus menjadikannya rujukan dalam penerapan teknik pembelajaran mesin untuk analisis sentimen di Indonesia, serta membantu pemerintah memahami reaksi publik terhadap kebijakan sektor energi nasional sebagai bahan pertimbangan dalam pengambilan kebijakan publik yang lebih baik.

1.4 Manfaat Penelitian

Manfaat yang diharapkan dari penelitian ini adalah sebagai berikut:

a. Manfaat Teoritis

Penelitian ini diharapkan dapat menambah pengetahuan dan pemahaman di bidang analisis sentimen, khususnya penerapan teknik machine learning menggunakan algoritma Logistic Regression pada data dari platform media sosial X. Selain itu, penelitian ini juga diharapkan bisa menjadi rujukan untuk penelitian berikutnya mengenai analisis sentimen opini publik di Indonesia.

b. Manfaat Praktis

Penelitian ini diharapkan memberikan gambaran umum mengenai sentimen masyarakat terhadap kebijakan kenaikan harga bahan bakar yang terbagi menjadi kategori positif, negatif, dan netral. Selain itu, hasil penelitian ini diharapkan dapat memberi wawasan bagi pemerintah dalam memahami tanggapan masyarakat terhadap kebijakan sektor energi nasional, sekaligus mendukung pengambilan kebijakan publik yang lebih terarah.

c. Manfaat Akademis

Penelitian ini diharapkan dapat menjadi rujukan bagi mahasiswa maupun peneliti lain yang ingin mengembangkan penelitian seputar analisis sentimen, text mining, Natural Language Processing (NLP), dan penggunaan algoritma Logistic Regression untuk klasifikasi data teks.