

Sentiment Analysis to Measure Public Trust in the Government Due to the Increase in Fuel Prices Using Naive Bayes and Support Vector Machine

Zakaria^{1*}, Kusri², Dhani Ariatmanto³

^{1,2,3}Informatics Department, Universitas Amikom Yogyakarta, Indonesia

¹zakaria@students.amikom.ac.id*; ²kusri@amikom.ac.id; ³dhaniari@amikom.ac.id

*corresponding author

ABSTRACT

The study examines public sentiment on the government's fuel price policy using an experimental approach and Twitter data obtained through API scraping. It applies sentiment analysis methods like Naïve Bayes, SVM, and Majority Voting. SVM achieved 85% accuracy, excelling in identifying negative sentiments, while Majority Voting reached 70% by considering confidence levels. Naïve Bayes struggled with neutral sentiments. They are combining methods to enhance the understanding of public sentiments on fuel price changes. The study highlights sentiment analysis' effectiveness in gauging reactions to fuel policies, with SVM offering more profound insights into sentiments related to fuel price hikes. Challenges remain in identifying neutral sentiments due to social media text brevity. These findings underscore the contextual importance of interpreting sentiment analysis. Leveraging these insights, governments can understand public perceptions better and devise improved communication strategies for sensitive economic policies like fuel price hikes, fostering better government-citizen interactions. The study aims to guide stakeholders in comprehending public perspectives within public policy, emphasizing the relevance of sentiment analysis for policy evaluation.

Keywords: Sentiment Analysis, Fuel Prices, Majority Voting, Naïve Bayes, Support Vector Machine

This is an open-access article under the [CC-BY-SA](#) license.



Article History

Received : Oct, 19th 2023

Accepted : Nov, 16th 2023

Published : Nov, 24th 2023

I. INTRODUCTION

Fuel has become an integral part of various aspects of life, fulfilling the needs of individuals and organizations and serving as a major contributor to Indonesia's oil and gas industry. Fuel not only plays a crucial role in supporting various sectors of life but also makes a significant contribution to the country's economy and serves as a source of revenue for the state budget (APBN) [1], [2]. It is important to note that the government plays a crucial role in regulating the pricing policies of Fuel (BBM), including subsidies such as *Pertalite*, *Solar*, and non-subsidized types like *Pertamax*[3].

Increases in fuel prices, whether subsidized or non-subsidized, have a widespread impact on society, affecting unemployment rates, economic growth (Gross Domestic Product/GDP), and inflation rates [4], [5]. Several factors contributing to the rise in fuel prices include rising production costs, which, in turn, affect consumer prices[6]. Another factor that influences this is inflation, which affects the cost of raw materials and labor, as well as government policies such as tax hikes or subsidy reductions[7], which can be observed in several countries, including Turkey and Indonesia[8][9]. In the context of fuel price hikes and government responses, sentiment analysis plays a crucial role[10]. Through sentiment analysis, we can understand the public's reactions and attitudes toward government policies[11], identify positive, negative, or neutral sentiments[12], and grasp the issues and needs that need to be addressed[13]. Research conducted by [14] highlights the importance of sentiment analysis in understanding public views on the government's policy regarding the Omnibus Law on the social media platform Twitter. In this study, tweet data related to the Omnibus Law policy was collected and processed through data pre-processing stages. The results of sentiment analysis reveal dominant sentiment patterns on Twitter regarding the Omnibus Law, aiding in understanding the public's attitudes and perceptions towards the policy. Despite some challenges, such as noisy data and limitations in generalizing results, this research emphasizes the significance of sentiment analysis in comprehending public responses to government policies on the Twitter social media platform.

The sentiment analysis was used in subsequent research [15] to determine the thoughts and perceptions of the general people regarding the Eid homecoming vaccine booster program. This study employed eight classification models: Naïve Bayes, Support Vector Machine (SVM), Decision Tree, Logistic Regression, Random Forest, K-Nearest Neighbor, AdaBoost, and XGBoost. The most effective classification model was SVM, attaining the highest accuracy score of 88% and an F1 score of 88%. Subsequently,

this SVM model was utilized to predict sentiments from 30,582 tweet data collected between March 22nd and May 2nd, 2022. The results indicated that 11,507 tweets conveyed negative sentiments (37.63%) while 19,075 conveyed positive sentiments (62.37%). These findings suggest that the government's policy in accelerating the COVID-19 booster vaccination program was well-received by making it a requirement for the Eid homecoming.[16] conducted sentiment analysis on Jokowi's policy regarding alcohol investment in Indonesia using Twitter API. The study dealt with 6,963 tweets before and after the policy revocation, with 963 manually annotated. Employing NLP & SVM techniques, the study achieved 66% accuracy. The results indicated a strong rejection of the initial policy. The revocation was considered a fitting step that effectively reduced negative comments on Twitter. The study also analyzed related keywords on Twitter. Previously, many tweets opposed the policy and requested its revocation. However, afterward, comments became more critical of the President's stance.

Sentiment analysis methods are used to measure and evaluate the public's opinions, views, and emotions regarding specific topics[17]. This approach involves collecting, processing, analyzing, and understanding sentiments and opinions text[18]. In the context of fuel price hikes, sentiment analysis becomes an effective tool for exploring public perspectives and attitudes toward such policies through social media. Several previous studies, such as those conducted by [19][20], have demonstrated the effectiveness of various sentiment analysis methods, including multimodal approaches. However, this study will focus on combining the Naive Bayes Classifier and SVM methods, along with the majority voting method, to enhance the accuracy and relevance of sentiment analysis in the context of public policies related to fuel price increases.

II. METHOD

This research adopts an experimental approach by implementing a single-method classification of naïve Bayes and SVM, combining them through majority voting. The descriptive nature of this study lies in the endeavor to draw conclusions regarding public perception of the government and to measure their confidence level after the increase in fuel prices. The approach taken in this research is quantitative, where quantifiable data is utilized to evaluate the performance of classification models, such as accuracy, precision, and recall. Based on the previous description, it can be determined the research flow in Figure 1

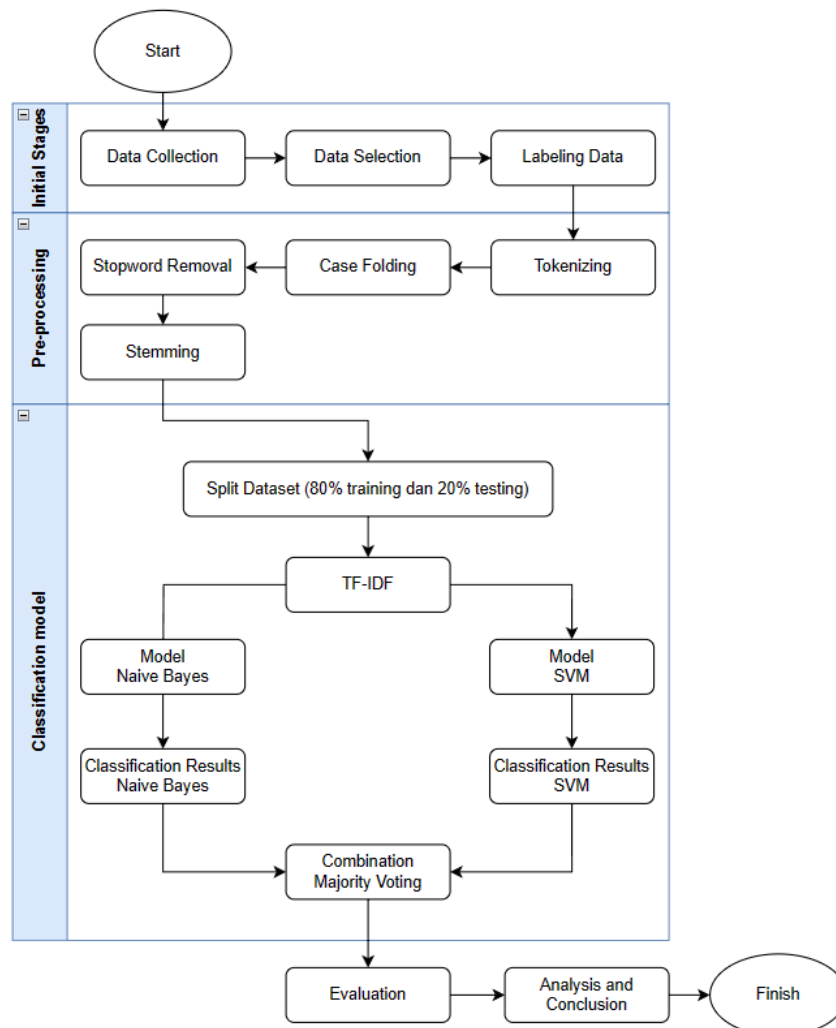


Figure 1. Research Flow

Figure 1 illustrates the research workflow encompassing a series of steps designed to combine Naive Bayes and SVM using majority voting in sentiment analysis on a dataset. These steps involve dataset preparation, pre-processing, feature extraction, Multinomial Naive Bayes and SVM classification implementation, and utilizing majority voting to merge prediction outcomes.

A. Data Collection

Data is obtained through scraping through the Twitter API using the keywords "BBM Mahal", "Bahan Bakar Naik", and "BBM Subsidi". The date range for data collection is from January 1st to May 31st, 2023. In this study, Twitter data retrieval using the Anaconda Application by implementing Tweepy in Python programming language was used to collect tweets related to the Fuel Oil (BBM) price increase policy, allowing researchers to easily and structurally access social data that can be applied in sentiment analysis related to public policy. The data collection flow can be seen in Figure 2.

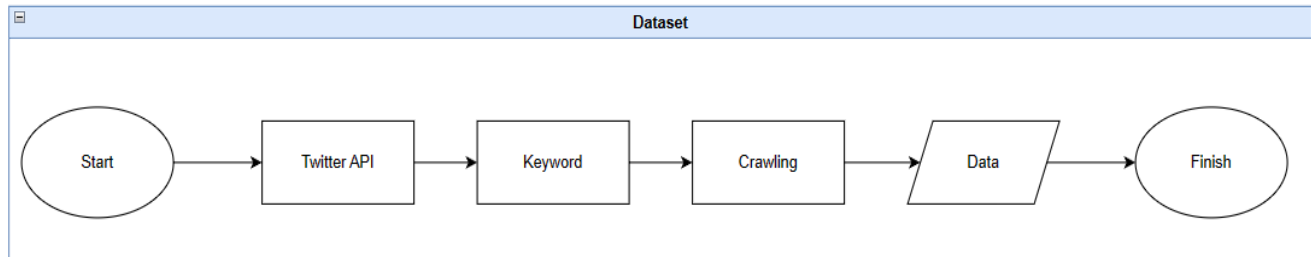


Figure 2. Stages of Data Collection

Figure 2 illustrates the data retrieval process, commencing with data collection through the Twitter API using Tweepy, a Python library that facilitates interaction with the Twitter API. Subsequently, relevant keywords related to the research topic are identified. These keywords are utilized in Twitter searches to gather the most pertinent data. The data collection is initiated by establishing a connection to the Twitter API using acquired access keys. Searches are conducted based on the predefined keywords. Following the retrieval of search results, Twitter data is extracted and stored for further analysis. This process may involve storing data in a specific format, such as a CSV file or within a database.

B. Data Selection

The dataset obtained through data collection will be initially screened to obtain data that genuinely represents public opinions. Tweets that violate copyright, irrelevant content, and data not aligning with the research context will be removed.

C. Text Pre-processing

The text pre-processing stages include tokenization, case folding, removal of common words (stopword removal), and stemming using the stemming algorithm from *Sastrawi*. Here is a detailed explanation for each stage, as shown in Figure 3.

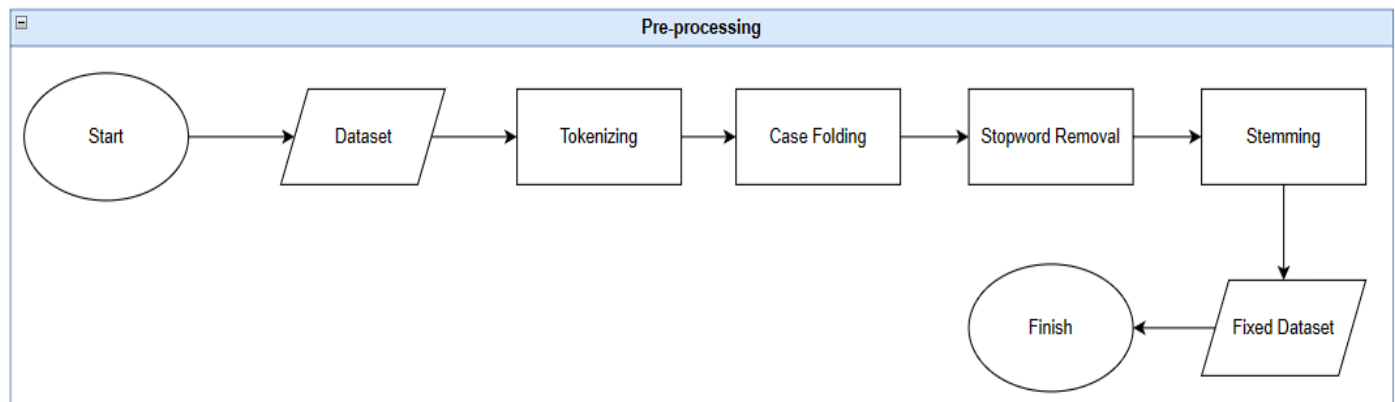


Figure 3. Data Pre-processing Stages

Figure 3 illustrates the crucial part of sentiment analysis: the pre-processing stage. This step is pivotal in preparing the data before implementing classification[21][22]. The focus is addressing potential errors in extracting features or attributes that can significantly impact sentiment analysis performance. During this stage, it is necessary to conduct data selection for processing in each document. Text pre-processing in sentiment analysis involves a series of steps without diminishing the substance of the data. Words with significant meaning are retained, ensuring the input data maintains its original meaning.

Additionally, attention must be given to handling colloquial language in Bahasa Indonesia, as each word in the sentiment analysis process becomes a feature in the dataset. Generally, pre-processing steps include lowercase adjustment (case folding), tokenization, removal of conjunction words (stop word removal), and stemming [23]. A detailed explanation of these steps is as follows:

- Tokenizing: This stage separates the text into individual words using spaces as separators. Each word becomes a separate unit ready for further processing.
- Case folding: After the cleaning stage, case folding is performed to convert all letters in the text to lowercase. This ensures no differences in handling words with capital or non-capital letters.
- Filtering or stopwords removal: In this stage, words considered to have no significant meaning in sentiment analysis will be removed. These words are often referred to as stopwords, such as conjunctions like "dan," (and) "atau" (or), and other common words. The goal is to reduce the excess of words that do not contribute to sentiment analysis.
- Stemming: The final stage in pre-processing uses the Sastrawi algorithm, which aims to transform words with affixes into their base form. For example, the word "membeli" will be transformed into "beli," the word "berlibur" will be transformed into "libur," and so on. This process helps reduce variations of different word forms in sentiment analysis and strengthens the representation of their underlying base words.

D. Sentiment Classification

The sentiment classification methods used include the Naive Bayes Classifier, Support Vector Machine (SVM), and combining these methods using the majority voting method. The formula gives the sentiment classification process using the Naive Bayes Classifier (NBC):

The process of classification of sentiment using Naïve Bayes is given by Equation (1).

$$P(\text{Sentimen}|\text{Text}) = \frac{P(\text{Text}|\text{Sentiment}) \times P(\text{Sentiment})}{P(\text{Text})} \quad (1)$$

The process of classification of sentiment using SVM is given by Equation (2).

$$y = \text{sign} \left(\sum_{i=1}^n \alpha_i y_{ii} K(x_i, x) \right) + b \quad (2)$$

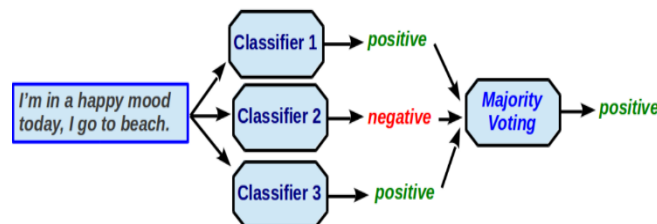


Figure 4. An example of Majority Voting as a combination rule

As for the majority voting method, the sentiment classification results are determined based on the majority of the classification results from both NBC and SVM. As seen in Figure 4 shows an example of a combination of rules.

E. Evaluation

In this phase, evaluation is conducted utilizing the Confusion Matrix, a technique employed to assess the performance of a classification model by comparing the predicted outcomes with the actual labels in the test data. The evaluation output encompasses performance metrics such as accuracy, precision, and recall using using Equation (3), (4), and (5), respectively.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

$$\text{Precision} = \frac{TP}{TP + FN} \quad (4)$$

$$\text{Recall} = \frac{TP}{TP + FP} \quad (5)$$

III. RESULT AND DISCUSSION

In this section, the researchers will explain the results obtained from the research. After applying the techniques outlined in the methodology section, these results are presented as raw data or outcomes. The data used in this study is a collection of text from the Twitter social media platform. This data was obtained through a data scraping process using the keywords "BBM Mahal," "Bahan Bakar Naik," and "BBM Subsidi". The data collection period ranges from January 1st to May 31st, 2023. Results can be seen in Figure 5.

	created_at	id_str	full_text	quote_count	reply_count	retweet_count	favorite_count	lang	user_id_str	conversation_id_str	username	tweet_url
0	Mon Feb 27 23:59:54 +0000 2023	1,63E+18	@msaid_didu Tambah 1 lg pak,BBM dan hrg² yg la...	0	1	0	0	in	1,45E+18	1,63E+18	Barry_OhBarry	https://twitter.com/Barry_OhBarry/status/16303...
1	Mon Feb 27 23:07:41 +0000 2023	1,63E+18	@VIVAcoid Iya betul berjuang buat wong cilik s...	0	0	0	0	in	1,51E+18	1,63E+18	dsanakrantau	https://twitter.com/dsanakrantau/status/163034...
2	Mon Feb 27 21:36:23 +0000 2023	1,63E+18	Luar bissa bahan minyak jelanta tetsebut sanga...	0	0	0	0	in	1,24E+18	1,63E+18	sam_mambela	https://twitter.com/sam_mambela/status/1630320...
3	Mon Feb 27 18:20:56 +0000 2023	1,63E+18	@sk3ltah @otongkoil nah kalo gitu gimana? tany...	0	1	0	0	in	38628241	1,63E+18	MillhamP	https://twitter.com/MillhamP/status/16302717827...
4	Mon Feb 27 18:14:40 +0000 2023	1,63E+18	@sk3ltah @otongkoil contoh ngga, harga BBM nai...	0	1	0	0	in	38628241	1,63E+18	MillhamP	https://twitter.com/MillhamP/status/16302702062...
...
7558	Thu Aug 31 12:22:00 +0000 2023	1,70E+18	@kegblgnunfaedh BBM yang bisa dijual di Indone...	0	0	0	0	in	1,66E+18	1,70E+18	adithidoth	https://twitter.com/adithidoth/status/16972232...
7559	Thu Aug 31 12:17:21 +0000 2023	1,70E+18	Kilas Balik Kisruh di Rezim ini mengenai Roadm...	0	0	0	1	in	7,22E+17	1,70E+18	OmJ_JeNggot	https://twitter.com/OmJ_JeNggot/status/1697222...
7560	Thu Aug 31 11:14:13 +0000 2023	1,70E+18	@tempodotco Rezim jkw ini kelihatan nya Rezim ...	0	0	0	0	in	1,17E+18	1,70E+18	RickyDiningrat1	https://twitter.com/RickyDiningrat1/status/169...
7561	Thu Aug 31 09:56:12 +0000 2023	1,70E+18	membantu meningkatkan perekonomian masyarakat ...	0	0	0	0	in	1138467176	1,70E+18	KementerianESDM	https://twitter.com/KementerianESDM/status/169...
7562	Thu Aug 31 09:56:11 +0000 2023	1,70E+18	ENERGI BERKEADILAN Halo sobat energi! Kenala...	0	2	1	5	in	1138467176	1,70E+18	KementerianESDM	https://twitter.com/KementerianESDM/status/169...

7563 rows x 12 columns

Figure 5.Scraping results

In Figure 5, the next step is to label a portion of the data after the scraping process collects and forms the dataset. This labeling is done to classify sentiments into three categories: negative, neutral, and positive. The labeling process is done manually, and the results are divided into two parts, namely, 80% training data and 20% testing data after pre-processing. Pre-processing is the initial stage in text data processing. This study's pre-processing steps include converting to lowercase, tokenizing, removing common words (stopword removal), and stemming using the *Sastrawi* stemming algorithm. The results of this process can be seen in Table I.

TABLE I
PRE-PROCESSING RESULTS

Text	ORIGINAL TEXT	After Pre-processing
1	@msaid_didu Tambah 1 lg pak,BBM dan hrg² yg lain tdk gampang naikðŸ˜f	['msaiddidu', 'tambah', 'sedang', 'pakbbm', 'harga', 'gampang', 'naik']
2	@VIVAcoid Iya betul berjuang buat wong cilik salah satunya. BBM naik di dukung.. UUD ciptakerja yang merugikan pekerja di di dukung.. Hukum tajam ke lawan tumpul ke kawan.. Ada yang mau nambahin?	['vivacoid', 'iya', 'betul', 'juang', 'buat', 'wong', 'cilik', 'salah', 'satu', 'bbm', 'naik', 'dukung', 'uud', 'ciptakerja', 'rugi', 'kerja', 'dukung', 'hukum', 'tajam', 'lawan', 'tumpul', 'kawan', 'mau', 'nambahin']
3	Luar bissa bahan minyak jelanta tetsebut sangat banyak. Semoga penemuan tersebut dapat menekan turun harga BBM yg terus naik melonjak	['luar', 'bissa', 'bahan', 'minyak', 'jelanta', 'tetsebut', 'sangat', 'banyak', 'moga', 'temu', 'sebut', 'tekan', 'turun', 'harga', 'bbm', 'terus', 'naik', 'lonjak']

The next step involves displaying a Word Cloud, a visual representation of text that can be seen in the image where the most frequently occurring words are depicted in Figure 6.



Figure 6. Word Cloud of Data Reviews

Figure 6 provides a quick visual overview of the relative frequency of words in the text using different sizes and colors, enabling the audience to instantly identify the most common or significant words in the text corpus. This word cloud not only displays information about the most frequently appearing words but also gives a general idea of the main topic or focus of the text. This visual method is frequently employed in text analysis, research, and data presentations to highlight important patterns and trends in the text in an engaging and easily understandable manner.

The next step involves data weighting using the TF-IDF (Term Frequency-Inverse Document Frequency) method. TF-IDF assigns weights to words in the text based on the frequency of word occurrence in a document (TF) and the inverse frequency of word occurrence across all documents (IDF). This approach allows emphasis on words that have a high impact within a document but are rare across the corpus. The following is the representation of data that has been weighted using TF-IDF, as depicted in Figure 7.

(0, 1245)	0.04846284129780544
(0, 4494)	0.08066848438233787
(0, 4552)	0.2972921802375916
(0, 7540)	0.3791298227979236
(0, 10000)	0.3004088820744835
(0, 10218)	0.30743454689027405
(0, 10781)	0.06693508535618317
(0, 11723)	0.3791298227979236
(0, 12972)	0.3037751765065919
(0, 15155)	0.31587399958316303
(0, 15419)	0.3791298227979236
(0, 15860)	0.3004088820744835
(1, 544)	0.18839925567762258
(1, 950)	0.08814029115851396
(1, 967)	0.3118075769917524
(1, 990)	0.08955147258532509
(1, 1206)	0.3025468871294005
(1, 2070)	0.13610634302036598
(1, 2178)	0.3025468871294005
(1, 2560)	0.17631719903856707
(1, 3152)	0.2889390671374106
(1, 4349)	0.1254937336902535
(1, 6960)	0.23724017022722227
(1, 7487)	0.14732811182055094
(1, 8319)	0.2908196350370132
...	
(6049, 13403)	0.39494152937894905
(6049, 14423)	0.39494152937894905
(6049, 14583)	0.24667291685832057
(6049, 16274)	0.2454813311265202

Figure 7. TF-IDF Weighting

We divided the dataset into training and testing subsets in our method testing. Typically, researchers use an 80:20 split, where 80% of the data is used for training, and 20% is used for testing. This division allows us to train the model with sufficient data and assess its performance with data it has never seen before. Classification testing was performed using Naïve Bayes, SVM, and Majority Voting. The accuracy test results can be seen in Table II.

TABLE II THE PERFORMANCE MODELS	
Method	Accuracy
Naïve Bayes	70%
SVM	85%
Majority Voting	70%

Table II The Naive Bayes method yielded an accuracy of approximately 70%. However, the performance of Naive Bayes tends to be lower than SVM's. This might be due to the strong (naive) assumption of Naive Bayes regarding feature independence, which may not be fully satisfied in complex social media text data. The Recall value for neutral sentiment is 0, indicating that Naive Bayes cannot identify neutral sentiment in this dataset. The Sentiment Classification Results Using the SVM Method yield an accuracy rate of approximately 85%. This indicates that SVM is quite effective in classifying public sentiment regarding the policy of fuel price increases. The classification results are also shown through Precision, Recall, and F1-score for each sentiment (negative, neutral, and positive) in Figure 8.

Classification Report:				
	precision	recall	f1-score	support
negative	0.86	0.93	0.90	880
neutral	0.75	0.04	0.07	77
positive	0.82	0.83	0.83	556
accuracy			0.85	1513
macro avg	0.81	0.60	0.60	1513
weighted avg	0.84	0.85	0.83	1513

Confusion Matrix:	
[[820 0 60]	
[35 3 39]	
[94 1 461]]	

Accuracy Score:	
0.8486450760079313	

Figure 8. Classification Results Using SVM

The Support Vector Machine (SVM) identifies negative and positive sentiments well, displaying relatively high Precision and Recall values. However, for neutral sentiments, its recall is notably low. This could be attributed to the inherent difficulty in identifying neutral sentiments from short texts commonly found in social media data. The Majority Voting Classification Results, which combine the outcomes of Naive Bayes and SVM, yield an accuracy of approximately 70%, as shown in Table II, mirroring the Naive Bayes results. This suggests that, in this particular case, employing majority voting does not enhance classification performance compared to using a single SVM. This circumstance might be due to Naive Bayes exhibiting lower performance, making a smaller contribution to the majority voting process. It is important to provide context regarding what each element represents and how each evaluation metric offers additional insights into the model's performance within the specific context observed in Figure 9.

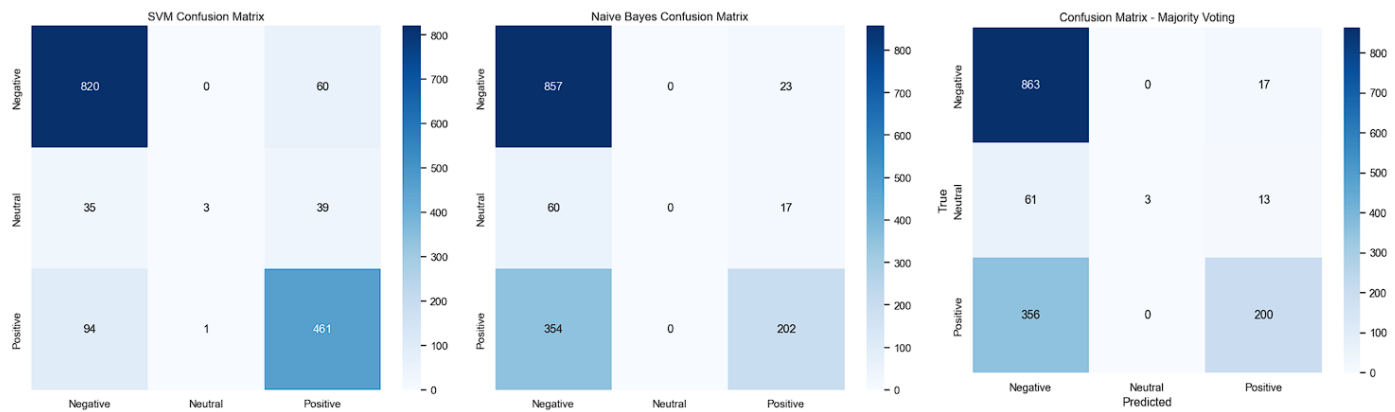


Figure 9. Confusion Matrix

Figure 9 illustrates that the Support Vector Machine (SVM) model excels in its ability to predict with good accuracy for the "negative" and "positive" classes, with high precision rates of 0.86 and 0.82, respectively. Furthermore, SVM also demonstrates solid recall for the "negative" (0.93) and "positive" (0.83) classes, indicating the model's capability to identify samples that truly belong to these classes accurately. However, it is important to note that when predicting the "neutral" class, SVM exhibits significantly lower performance with a recall of 0.04, resulting in a low F1-score (0.07) for this class.

On the other hand, the Naive Bayes model displays low accuracy in predicting all classes, particularly with an extremely low precision rate for the "neutral" class (0.00), signifying the model's inability to predict this class correctly. Although it achieves high recall for the "negative" class (0.97), the performance of Naive Bayes is impacted in measuring the F1-score, especially for the "negative" (0.80) and "positive" (0.51) classes.

Meanwhile, the Majority Voting model better predicts the "neutral" class with extremely high recall and precision rates (1.00). However, this model still faces challenges in predicting the "positive" class with a relatively low recall of 0.36. Majority Voting demonstrates a more balanced performance between the "neutral" and "positive" classes compared to SVM and Naive Bayes. Regarding overall accuracy, SVM approaches a value of 0.85, while Naive Bayes and Majority Voting have accuracies around 0.70. Despite SVM displaying better overall performance, Majority Voting and Naive Bayes have their respective strengths and weaknesses in predicting specific classes. This may be because Naive Bayes has lower performance, thus making a smaller contribution in the majority voting. The confidence levels (probabilities) for each sentiment of the public toward government policies can be seen in Table III.

TABLE III

THE CONFIDENCE LEVEL (PROBABILITY) FOR EACH SENTIMENT

SENTIMENT	Average probability
Negative	64.65%
Neutral	2.98%
Positive	32.38%

The results of Table III above illustrate that the majority of sentiment expressed by the public regarding the fuel price increase policy is negative, followed by positive sentiment and neutral sentiment, which is quite low. This confidence level provides further insight into the distribution of public sentiment toward the policy. The results of this research provide important insights into the sentiment analysis of the public regarding the fuel price hike policy.

Sentiment classification methods such as Support Vector Machine (SVM) and Naive Bayes are crucial in analyzing these sentiments. The results show that SVM outperforms Naive Bayes in classifying sentiments related to the fuel price hike policy. SVM achieves an accuracy of approximately 85%, while Naive Bayes is around 70%. This may be attributed to SVM's ability to handle complex relationships among features in social media text data. SVM is more effective in recognizing negative and positive sentiments, which is essential in public policy.

However, challenges were also encountered in identifying neutral sentiment. This is primarily related to short texts commonly found on social media that are difficult to identify accurately. The results show very low Recall values for neutral sentiment in all methods, approaching zero. This reflects the difficulty in distinguishing neutral sentiments from informative or descriptive text. This challenge underscores the importance of considering context and data type when interpreting sentiment analysis results.

Furthermore, using the majority voting method to combine the results of SVM and Naive Bayes did not significantly improve classification performance compared to a single SVM. This indicates that Naive Bayes contributed less to the final results in this case. Combining methods may yield better results if each method contributes in a balanced or complementary manner.

The level of public confidence in sentiments is also a crucial aspect of this research. Negative sentiment has an average probability of around 64.65%, indicating significant concerns in the public regarding the fuel price hike policy. On the other hand, positive sentiment, with an average probability of around 32.38%, suggests that a portion of the population understands or supports the policy. Neutral sentiment has a very low probability (around 2.98%), indicating that the public tends to have strong positive or negative opinions about the fuel price hike policy.

The implications in the context of public policy are highly significant. The government can use these results as a guide to understanding the perception and reactions of the public to sensitive economic policies like fuel price hikes. Understanding public sentiment is the first step in formulating more effective and acceptable policies. Considering the public's confidence level, the government can make more appropriate efforts to communicate and respond to these policies, build public trust, and foster more positive interactions between the government and citizens.

IV. CONCLUSION

This research applies sentiment analysis methods, namely Naïve Bayes, Support Vector Machine (SVM), and Majority Voting, to evaluate public perspectives on the government's fuel price policy using data from the Twitter platform. The study results indicate that SVM achieved the highest accuracy rate of about 85%, particularly excelling in identifying negative sentiments. However, Naïve Bayes exhibited lower performance in recognizing neutral sentiments. Combining methods through Majority Voting showed the potential to deepen the understanding of public sentiments regarding fuel price changes. The research underscores the effectiveness of sentiment analysis in comprehending reactions to fuel policies, especially with SVM providing more profound insights into fuel price hikes.

Nonetheless, challenges persist in identifying neutral sentiments due to the limitations of short text in social media. The implications of these findings highlight the importance of considering context and data types when interpreting sentiment analysis results. These research outcomes have a significant impact in the context of public policy. Governments can leverage these insights to understand public perceptions better, formulate more effective communication strategies concerning sensitive policies like fuel price hikes, and enhance interactions between the government and citizens.

REFERENCES

- [1] G. Rozy Hrp, N. Aslami, and P. Studi Manajemen Fakultas Ekonomi Bisnis Islam, "Analisis Damfak Kebijakan Perubahan Publik Harga BBM terhadap Perekonomian Rakyat Indonesia," *Jurnal Ilmu Komputer, Ekonomi dan Manajemen*, vol. 2, no. 1, pp. 1464–1474, 2022.
- [2] D. Kompensasi Kenaikan Harga Bahan Bakar Minyak Terhadap, D. Lauradina, and A. Murwiati, "Pengeluaran Rumah Tangga Miskin Dengan Metode Difference in Difference (Studi Kasus: Kecamatan Panjang Kota Bandar Lampung),"
- [3] T. D. Arista, M. Fikry, L. Oktavia, U. Islam, N. Sultan, and S. Kasim, "Klasifikasi Sentimen Masyarakat di Twitter terhadap Kenaikan Harga BBM dengan Metode K-NN," vol. 5, pp. 140–150, 2023.
- [4] A. Tanza, R. P. Maulidya, T. P. Junita, and E. Widodo, "Analisis Pengaruh Kenaikan Harga BBM Terhadap Pergerakan Saham Sektor Transportasi dan Logistik," *Dialektika : Jurnal Ekonomi dan Ilmu Sosial*, vol. 8, no. 1, pp. 52–62, 2023, doi: 10.36636/dialektika.v8i1.2044.

- [5] W. Wardani, S. Suriana, S. U. Arfah, Z. Zulaili, and P. S. Lubis, "Dampak kenaikan Bahan Bakar Minyak (BBM) Terhadap Inflasi dan Implikasinya Terhadap Makroekonomi di Indonesia," *All Fields of Science Journal Liaison Academia and Society*, vol. 2, no. 3, pp. 63–70, 2022, doi: 10.58939/afosj-las.v2i3.338.
- [6] M. I. Al Islami, Q. Mauludin, S. D. Claritsa, E. Permana, and R. Amyulianthy, "Pengaruh Harga Bbm Terhadap Daya Beli Masyarakat di Indonesia, Malaysia dan Singapura," *Jurnal Ilmu Komputer dan Bisnis*, vol. 14, no. 1, pp. 180–187, May 2023, doi: 10.47927/jikb.v14i1.451.
- [7] R. Harun, R. Ishak, and S. S. Panna, "Analisis Sentimen Opini Publik Pengguna Twitter Terhadap Kenaikan Harga BBM Menggunakan Algoritma Naïve Bayes," *Copyright @BALOK*, vol. 2, no. 1, 2023.
- [8] Sarbaini and Nazaruddin, "Pengaruh Kenaikan BBM Terhadap Laju Inflasi di Indonesia," *Jurnal Teknologi dan Manajemen Industri Terapan (JTMIT)*, vol. 2, no. 1, pp. 25–32, 2023.
- [9] B. F. R. Al Fajrin, A. Z. Jinan, P. S. Auliyah, A. F. Wulandari, and N. Tambunan, "The Impact of the Drastic Increase in Fuel Oil (Bbm) Prices," *Journal of Social Research*, vol. 2, no. 3, pp. 640–645, 2023, doi: 10.55324/josr.v2i3.709.
- [10] F. Sahid and R. R. Fauzy, "Dampak Kenaikan Harga BBM terhadap RBP di Kontraktor X Karena Kelangkaan BBM di Pandemi Covid-19," *J-MAS (Jurnal Manajemen dan Sains)*, vol. 8, no. 2, p. 1298, Oct. 2023, doi: 10.33087/jmas.v8i2.1186.
- [11] S. Rukaiyah, "ANALISIS MANAJEMEN STRATEGI DALAM MENGHADAPI DAMPAK KENAIKAN BBM (STUDI KASUS NELAYAN DI KAB. MERAUKE)," *POMA JURNAL : PUBLISH OF MANAGEMENT*, vol. 1, pp. 70–83, 2023.
- [12] I. Karamitsos, S. Albarhami, and C. Apostolopoulos, "Tweet Sentiment Analysis (TSA) for Cloud Providers Using Classification Algorithms and Latent Semantic Analysis," *Journal of Data Analysis and Information Processing*, vol. 07, no. 04, pp. 276–294, 2019, doi: 10.4236/jdaip.2019.74016.
- [13] M. R. Nurhusen, J. Indra, and ..., "Analisis Sentimen Pengguna Twitter Terhadap Kenaikan Harga Bahan Bakar Minyak (BBM) Menggunakan Metode Logistic Regression," *Jurnal Media ...*, vol. 7, pp. 276–282, 2023, doi: 10.30865/mib.v7i1.5491.
- [14] E. A. Sukma, A. N. Hidayanto, A. I. Pandesenda, A. N. Yahya, P. Widharto, and U. Rahardja, "Sentiment Analysis of the New Indonesian Government Policy (Omnibus Law) on Social Media Twitter," *Proceedings - 2nd International Conference on Informatics, Multimedia, Cyber, and Information System, ICIMCIS 2020*, no. July 2017, pp. 153–158, 2020, doi: 10.1109/ICIMCIS51567.2020.9354287.
- [15] A. Pratama, R. I. Alhaqq, and Y. Ruldeviyani, "Sentiment Analysis of the Covid-19 Booster Vaccination Program As a Requirement for Homecoming During Eid Fitr in Indonesia," *J Theor Appl Inf Technol*, vol. 101, no. 1, pp. 248–261, 2023.
- [16] K. W. Lestari and T. Coan, "Evaluating Jokowi's Policy on Alcoholic Drink Investment in Indonesia by using sentiment analysis," *Journal of Advanced Research in Social Sciences*, vol. 5, no. 1, pp. 35–45, 2022, doi: 10.33422/jarss.v5i1.634.
- [17] L. Yue, W. Chen, X. Li, W. Zuo, and M. Yin, "A survey of sentiment analysis in social media," *Knowl Inf Syst*, vol. 60, no. 2, pp. 617–663, 2019, doi: 10.1007/s10115-018-1236-4.
- [18] A. Kukkar, R. Mohana, A. Sharma, A. Nayyar, and M. A. Shah, "Improving Sentiment Analysis in Social Media by Handling Lengthened Words," *IEEE Access*, vol. 11, no. January, pp. 9775–9788, 2023, doi: 10.1109/ACCESS.2023.3238366.
- [19] P. Zola, P. Cortez, C. Ragno, and E. Brentari, *Social Media Cross-Source and Cross-Domain Sentiment Classification*, vol. 18, no. 5, 2019. doi: 10.1142/S0219622019500305.
- [20] A. Khudhair, A. Khalil, H. A. Hussein, Q. Mohammed, and S. Alaa, "TWITTER SENTIMENT ANALYSIS USING AN ENSEMBLE MAJORITY VOTE CLASSIFIER," vol. 55, no. 1, pp. 1–7, 2020.
- [21] F. Handayani and S. Pribadi, "Implementasi Algoritma Naive Bayes Classifier dalam Pengklasifikasian Teks Otomatis Pengaduan dan Pelaporan Masyarakat melalui Layanan Call Center 110," *Jurnal Teknik Elektro*, vol. 7, no. 1, pp. 19–24, 2015.
- [22] C. Troussas, A. Krouska, and M. Virvou, "Evaluation of ensemble-based sentiment classifiers for Twitter data," *IISA 2016 - 7th International Conference on Information, Intelligence, Systems and Applications*, 2016, doi: 10.1109/IISA.2016.7785380.
- [23] I. M. A. Agastya, "Pengaruh Stemmer Bahasa Indonesia Terhadap Peforma Analisis Sentimen Terjemahan Ulasan Film," *Jurnal Tekno Kompak*, vol. 12, no. 1, p. 18, 2018, doi: 10.33365/jtk.v12i1.70.