InPrime: Indonesian Journal of Pure and Applied Mathematics
Vol. 7, No. 1 (2025), pp. 88 – 100, doi: 10.15408/inprime.v7i1.42717
p-ISSN 2686-5335, e-ISSN: 2716-2478

A Study on Sentiment Analysis of Public Response to The New Fuel Price Policy In 2022: A Support Vector Machine Approach

Niluh Putu Aprillia Puspitadewi Sudarsana Putri[1], Dwi Shinta Angreni[1], and I Wayan Sudarsana[2]*
[1]Department of Informatics, Faculty of Engineering, Tadulako University, Palu, Indonesia
[2]Department of Mathematics, Faculty of Mathematics and Natural Sciences, Tadulako University, Palu, Indonesia
Email: *sudarsana.iwayan@untad.ac.id

Abstract

The Indonesian government's decision to raise fuel prices in 2022, following a global surge in crude oil prices, triggered widespread public debate. Understanding public sentiment toward such policy decisions is essential for determining the appropriate timing of implementation while minimizing negative reactions. This study aims to classify public sentiment regarding the fuel price hike using the Support Vector Machine (SVM) algorithm. Data were collected from Twitter through web scraping using the SNScrape library in Python. A total of 3,000 tweets were gathered and underwent preprocessing steps such as case folding, tokenization, stopword removal, and stemming. The classification model was built in Google Colab using the SVM algorithm to categorize tweets as positive (+) or negative (–). Model performance was evaluated using a confusion matrix, achieving an accuracy of 81.0%. The results showed that 63.6% of public responses were negative, while 36.4% were positive. Additionally, it was observed that the accuracy converged to 81.1% as the number of training iterations increased. The findings were presented through word clouds and pie charts to enhance interpretability, and a simple graphical user interface (GUI) was developed for user interaction. The study indicates that the government's repeated delays in implementing the price adjustment may have reflected sensitivity to public sentiment. This research demonstrates the potential of sentiment classification as a tool for evidence-based policymaking, offering insights into the social dynamics surrounding policy changes. Future research could expand by incorporating multi-class sentiment categories or real-time data for dynamic policy evaluation.

Keywords: Fuel price; Public opinion; Sentiment analysis; Social media; SVM.

Abstrak

Keputusan pemerintah Indonesia untuk menaikkan harga bahan bakar minyak pada tahun 2022 dan disusul oleh lonjakan harga minyak mentah global, memicu perdebatan publik yang meluas. Memahami sentimen publik terhadap keputusan kebijakan tersebut sangat penting untuk menentukan waktu implementasi yang tepat untuk meminimalkan reaksi negatif. Penelitian ini bertujuan untuk mengklasifikasikan sentimen publik terhadap kenaikan harga bahan bakar minyak menggunakan algoritma Support Vector Machine (SVM). Data dikumpulkan dari Twitter melalui web scraping menggunakan pustaka SNScrape dalam bahasa Python. Sebanyak 3.000 tweet dikumpulkan dan dilakukan tahap praproses seperti case folding, tokenization, stopword removal, dan stemming. Model klasifikasi dibangun di Google Colab menggunakan algoritma SVM untuk mengkategorikan tweet sebagai positif (+) atau negatif (–). Kinerja model dievaluasi menggunakan matriks confusion dan mencapai akurasi 81,0%. Hasil penelitian menunjukkan bahwa 63,6% tanggapan publik bersifat negatif, sedangkan 36,4% bersifat positif. Selain itu, akurasi konvergen menjadi 81,1% seiring dengan peningkatan jumlah iterasi pelatihan. Temuan tersebut disajikan melalui word cloud dan diagram pai untuk meningkatkan interpretabilitas, dan graphical user interface (GUI) sederhana dikembangkan untuk interaksi pengguna. Studi ini menunjukkan bahwa penundaan berulang pemerintah dalam menerapkan penyesuaian harga mungkin mencerminkan kepekaan terhadap sentimen publik. Penelitian ini menunjukkan potensi klasifikasi sentimen sebagai alat untuk pembuatan kebijakan berbasis bukti, yang menawarkan wawasan tentang dinamika sosial seputar perubahan kebijakan. Penelitian di masa mendatang dapat diperluas dengan menggabungkan kategori sentimen multikelas atau data waktu nyata untuk evaluasi kebijakan yang dinamis.

Kata Kunci: *Bahan bakar; Opini public; Analisis sentiment; Mesia social; SVM.*

2020MSC: 62H30, 91D30.

*) Corresponding author
Submitted December 2nd, 2024, Revised May 30th, 2025,
Accepted for publication May 31st, 2025, Published Online May 31st, 2025

1. INTRODUCTION

Today, social media is increasingly integrated with the advancement of information and communication technology. Not only a few people use it in their daily activities. According to [1], 59% or 160 million Indonesians actively use social media. Twitter is one of the most widely used social media platforms in the community today. Twitter is a social media where users can express their writing in 280 characters. This makes Twitter a media for people to express opinions. In 2020, the number of daily active Twitter users reached 166 million, which increased by 24% from 134 million in 2019 [2]. As the user base grows, the volume of posted tweets naturally increases, covering a wide range of topics, including public opinion and comments on economics, social behavior, natural phenomena, trade, education, entertainment, and many other subjects [3][4][5]. Twitter is also a significant platform for discussing trending topics worldwide. One such trending topic at the end of 2022 was the Indonesian government's decision to raise fuel prices. This policy was a response to the global surge in oil prices driven by the conflict between Russia and Ukraine. The world oil prices, and the Indonesian Crude Price (ICP) reached US$95.45 per barrel, far exceeding the ICP assumption in the 2022 State Budget, which was only US$63 per barrel. This increase in fuel prices, combined with rising interest rates and liquidity, had a notable impact on developing countries like Indonesia, leading to challenges in managing existing subsidy budgets [6]. The policy adjustment aimed at raising fuel prices swiftly gained prominence on various social media platforms, including Twitter. In the aftermath of the policy's enactment, Twitter feeds were inundated with public responses to the announcement of the fuel price increase. The examination of various societal responses to the benefits and drawbacks that arise is a compelling subject of study. Several methodologies can be employed to analyze individuals' reactions, especially on social media, one of which is sentiment analysis. Sentiment analysis is a technique that quantifies and assesses specific cases or objects, enabling the derivation of conclusions and decisions based on text in the form of sentences or documents [7]. Several tools can be utilized in sentiment analysis, including the Support Vector Machine (SVM) algorithm standing out as one of the most effective options. Motivated by [2], [8] and [9], the SVM algorithm is renowned for its capacity to deliver the most accurate results compared to other classifier algorithms.

The research topic based on the Support Vector Machine (SVM) model is highly intriguing, encompassing various aspects of study in a field with broad applications [10]. For instance, [8] conducted facial emotion recognition using SVM and subsequently [1] confirmed the advantages of multiport Electric Vehicle (EV) power charging circuits using an SVM approach. In [3], focused on developing an accurate algorithm for data analysis and its application using the RBF SVM kernel method. Shao and Gao [11] conducted a study to develop a prediction model for assessing or estimating gas prominence levels using SVM. Serin et al. [12] performed gender classification based on fingerprint identification by combining automatic feature extraction with SVM. In [11], accurately predicted soil moisture content in tea plantations using an SVM-based model. Lin [5] detected pedestrians through image processing and object detection algorithms using SVM based on computer vision. Acharya et al. [13] devised an automatic machine learning-based disease detection method using SVM. Abdelfattah [14] used the SVM method to address multi-label topic classification issues in various datasets by converting multi-label datasets into a single label. The two-stage estimation process of water surface levels is a crucial aspect in control and water flow structure design, as [13] demonstrated that water surface profiles generated by simulation using the SVM model exhibit a high degree of conformity with empirical data. Predicting coal and gas outbursts with a low index (LI-CGO) can be challenging and pose a threat to efficient coal mining. [14] asserted that SVM prediction accuracy surpasses that of Back Propagation Neural Network and Distance Discriminant Analysis in predicting LI-CGO. Furthermore, according to [15], the SVM model excels in stock price prediction but may not be suitable for large datasets. Fan [16] conducted a study that combined texts and image datasets generated from social media to perform sentiment analysis using the Deep Multimodal Fusion Model (DMFM), including the SVM method.

This section reviews the results of the previous research about public sentiment related to certain topics with the implementation of the SVM Algorithm. Hendrastuty et al. [17] used SVM to analyze the sentiment of the Indonesian people towards the pre-employment card program as the government's effort to overcome unemployment and victims of layoffs of workers through the Twitter platform. This research uses a dataset of 2000 data by using the data search keyword "Pre-employment" will be classified into three types of sentiment, namely positive, negative and neutral. This research compares two kernels, namely linear and RBF. The evaluation results were carried out on linear kernel accuracy values of 98.67%, precision 98%, recall 99%, and F1-Score 98%, while RBF kernel accuracy values were 98.34%, precision 97%, recall 98%, F1-Score 98%. So, it can be concluded that public sentiment from Twitter users towards the pre-employment card program during the pandemic is more neutral at 98.34%. Based on the results of the evaluation carried out in this research, the accuracy value of the linear kernel produces an accuracy value of 98.34%.

Research related to public sentiment classification using the Support Vector Machine (SVM) algorithm has shown promising results. Wati and Ernawati [18], for example, compared the performance of linear kernels and Radial Basis Function (RBF) in classifying public sentiment towards the Community Activity Restriction (PPKM) policy in the Java and Bali regions. Using 1,514 tweet data (757 positive class tweets and 757 negative class tweets), they showed that the RBF kernel produced a higher accuracy of 98.67%, compared to the linear kernel, which reached 86%. These results strengthen the evidence that the selection of kernels in SVM greatly affects classification performance. Based on these findings and considering the importance of understanding public perception of government policies, this study proposes an analysis of public sentiment on Twitter towards the 2022 Fuel Price Increase using the SVM Algorithm. This study aims to identify and classify public opinion regarding the fuel price increase policy issued by the Indonesian government in 2022 using a machine learning approach. The main contribution of this study is to provide an SVM-based classification model that is able to provide a quantitative picture of public reactions on social media. The results of this study are expected to be a strategic consideration for the government in assessing public readiness before implementing similar policies in the future.

2. METHODS

The design of the research methodology in this study was conducted in accordance with the following flowchart. The process began by using the text preprocessing method to clean the datasets taken from Twitter. Before using the datasets for SVM modeling, split and transform them once they're ready. After the model is created, it needs to be tested and evaluated using the confusion matrix and K-fold cross-validation methods. The final outputs are a graphical user interface (GUI) and data visualization for the classification model.

2.1. Text Preprocessing

Text Preprocessing is the first step in data processing to select words in data. Where in this study, it was tweet data, so it can produce more concise words that contain sentiments by selecting and removing words that are not needed [17]. The steps taken in text preprocessing are case folding, where all characters in the document transform into lowercase. The next step, spell normalization, is repairing the shortened words. The tokenizing step is used to break sentences into words per word; after filtering/stopword removal used to eliminate words that are not important and have no meaning, such as conjunctions and pronouns. And the last thing to do is stemming, which is to find the base form in each word [18].

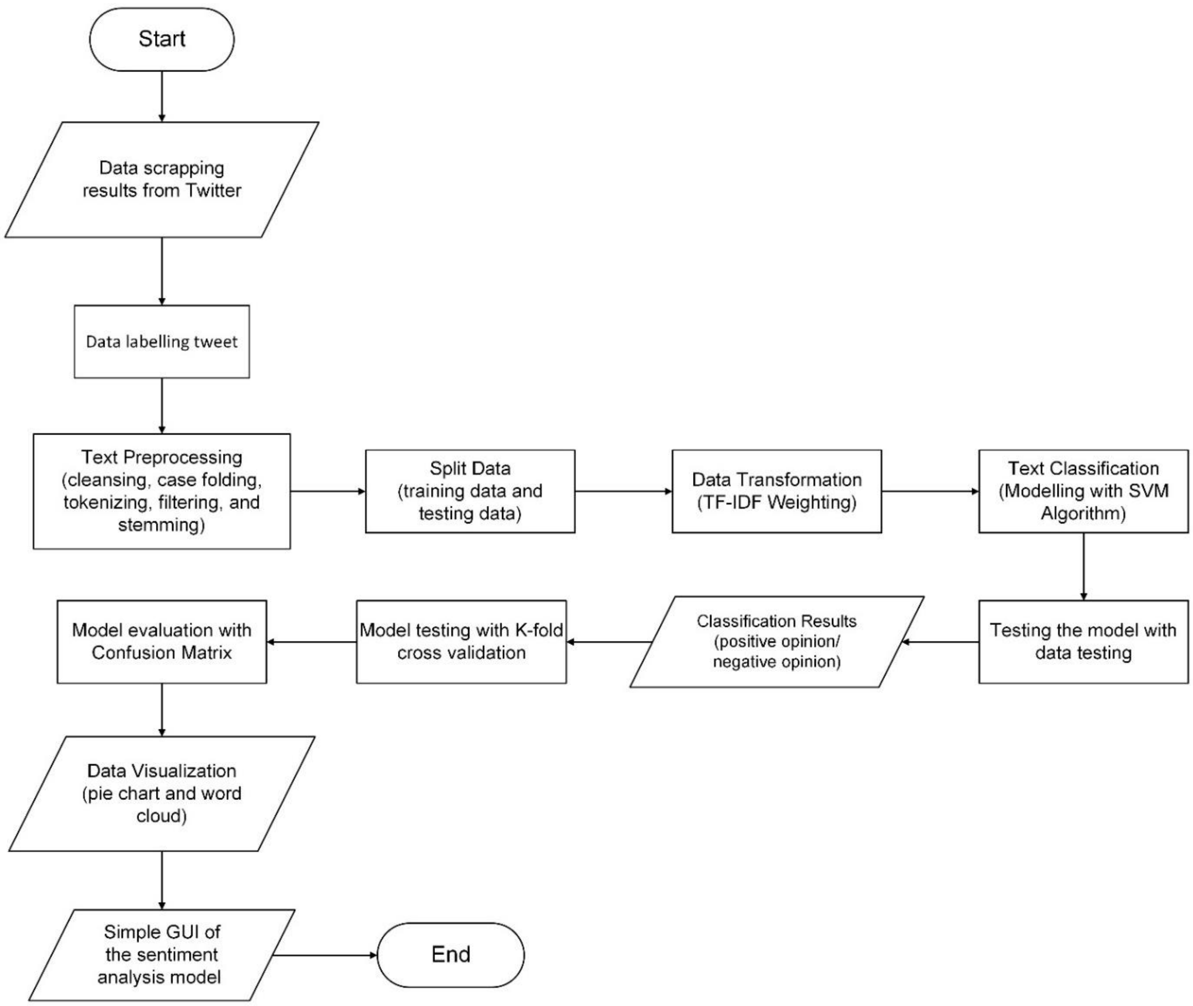

Figure 1. Research methodology design

2.2. Term Frequency-Inverse Document Frequency

The word weighting method used in documents is the Term Frequency-Inverse Document Frequency (TF-IDF), which is a method used to assign the weight of each word to determine how related the word is in a document [16][19]. This method combines two concepts, namely the concept of the frequency of occurrence of a term in a document and the inverse of the frequency of documents containing the word. This will be able to increase the proportion of the number of documents that can be found again, and which are considered relevant simultaneously. So, the term criteria to be obtained are terms that often appear in individual documents but are rarely found in other documents. The formula for TF-IDF is as follows:

$$tf = 0{,}5 + 0{,}5\left(\frac{tf}{max(tf)}\right), \tag{1}$$

$$idf_t = \log\left(\frac{D}{df_t}\right), \tag{2}$$

$$W_{d,t} = tf_{d,t} \times IDF_{d,t}, \tag{3}$$

where tf is the number of words searched for in a document, $max(tf)$ is the highest number of occurrences of the term on the same document, D is total documents, df_t is the number of documents containing term t, IDF is inversed document frequency, d is the d-th document, t is the t-th word of the keyword, and W is the weight of the d-th document to the t-th word.

2.3. Support Vector Machine (SVM) Algorithm

SVM is a learning system that uses a hypothetical space in the form of linear functions in a high-dimensional feature space and implements a learning bias derived from statistical learning theory, which is trained with learning algorithms. The theory underlying SVM has developed since the 1960s but was only introduced by Vapnik, Boser and Guyon in 1992. In simple terms, the SVM concept is an attempt to find the "best" hyperplane which plays an important role as the boundary line for two classes. SVM searches for these hyperplanes based on support vectors and margins. Support vectors are all data vectors that are closest to the hyperplane, while the margin represents the width of the separating hyperplane [20]. Support Vector Machine (SVM) includes machine learning (supervised learning), which can predict classes based on the results of the training process. By conducting training using input data in numerical form as the result of feature extraction, a pattern is obtained, which will later be used in the labelling process [21].

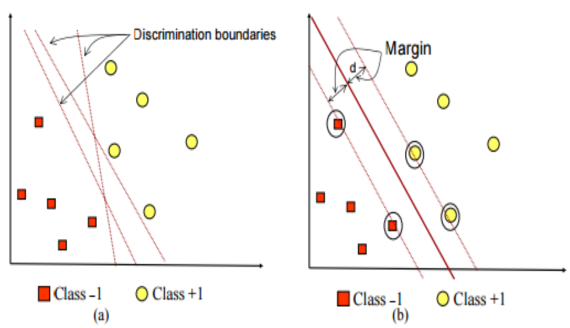

Figure 2. Best hyperparameter that separates positive and negative classes [21]

There are several cases where the classification problems found are not always linear, but there are also cases of non-linear classification. The kernel trick is a development of the SVM method to solve non-linear classification problems. To solve non-linear problems, SVM is modified by including Kernel functions. In non-linear SVM, data x is first mapped by the function Φ(x) to a higher dimensional vector space. In this new vector space, a hyperplane that separates the two classes can be constructed [22]. The commonly used kernel trick formula is as follows:

Polynomial Kernel: $$K(x, y) = (x, y + c)^d, \tag{4}$$
Sigmoid Kernel: $$K(x, y) = \tanh(\sigma(x, y) + c), \tag{5}$$

Radial Basis Function (RBF) Kernel: $$K(x, y) = e^{-\frac{\|x-y\|^2}{2\sigma^2}}, \quad (6)$$
where c is constant, d is degree, and e is natural number.

2.4. K-Fold Cross Validation

K-Fold Cross-validation is a validation method that aims to assess the performance of a model built by an algorithm by randomly dividing data samples (training data and test data) K times, where K is the size of the data partition number used for training and testing division. The data will be processed for several K experiments in each scenario using testing data and training data, respectively. K-fold cross-validation repeatedly divides data into training data and testing data, where each data could become testing data [23].

2.5. Confusion Matrix

The Confusion Matrix is a technique used to evaluate the classification model to estimate whether the object is true or false. A matrix of predictions will be compared with the original class, which contains actual information and predicted classification values. After the system successfully classifies tweets, it takes a measure to determine how valid or appropriate the classification has been made by the system [21]. Confusion matrix is a table that states the classification of the number of correct test data and the wrong number of test data [24].

3. RESULTS

3.1. Data collection

The dataset used in this research is from Twitter, especially from verified news accounts that post news related to the government's policy of increasing fuel prices. The news has been taken since September 3rd, 2022. After collecting news accounts, tweets are taken from the comments section in every post using the library SNScrape. The data that has been collected is 5076 data with source in Table 1, but the data will be filtered again, so it will be used only 2161 data.

This study adopts a case study approach, focusing on six influential Twitter accounts that represent key perspectives within the discourse on public response to the new fuel price policy in 2022. These accounts were selected based on their relevance, diversity of opinion, and engagement levels. While the small sample size limits generalizability, the purpose of this approach is to enable a deeper, more focused analysis of the arguments and narratives shaping the conversation. Insights from this case study can serve as a foundation for further research involving larger datasets or additional methodologies.

Table 1. Data sources

No	Twitter Accounts	Number of Data
1	@asumsico	1663
2	@catchmeupid	437
3	@cnnindonesia	2022
4	@collegemenfess	407
5	@convomf	245
6	@txtdrimedia	302
Total		**5076**

3.1. Preprocessing Data

The first step in text preprocessing is case folding, at which stage all alphabetic characters in the dataset are changed to lowercase. The results of case-folding can be seen in Figure 3.

```
Case Folding Result :

0              kaya akan minyak tapi minyaknya mahal
1             akhir dari era keliling cari makan malam
2    pejuang pulang pergi menangis benar ini bensin...
3    biasanya isi pertamax penuh 32 ribu sekarang j...
4    andai gaji naiknya serajin naiknya harga2 kebu...
```

Figure 3. Case folding result

The next stage is the cleansing process, where punctuation marks or characters other than the alphabet are eliminated, which will not be needed in making a classification model. The results of the data cleansing stages can be seen in Figure 4.

```
Cleansing Result :

0              kaya akan minyak tapi minyaknya mahal
1             akhir dari era keliling cari makan malam
2    pejuang pulang pergi menangis benar ini bensin...
3    biasanya isi pertamax penuh ribu sekarang jadi...
4    andai gaji naiknya serajin naiknya harga kebut...
```

Figure 4. Cleansing result

The next stage is tokenizing, where the data, which was originally in the form of a sentence, is separated into words or tokens, with a "," sign separating each token. The results of the tokenizing process can be seen in Figure 5.

```
Tokenizing Result :

0        [kaya, akan, minyak, tapi, minyaknya, mahal]
1     [akhir, dari, era, keliling, cari, makan, malam]
2    [pejuang, pulang, pergi, menangis, benar, ini,...
3    [biasanya, isi, pertamax, penuh, ribu, sekaran...
4    [andai, gaji, naiknya, serajin, naiknya, harga...
```

Figure 5. Tokenizing result

Stopword removal will be the next stage of this preprocessing step. At this stage, the data will be cleaned up with words that have no meaning, such as conjunctions, pronouns, and other words that do not affect any sentiment in the commentary text. The results of the stopword removal process can be seen in Figure 6.

```
0                     [kaya, minyak, minyaknya, mahal]
1                   [era, keliling, cari, makan, malam]
2    [pejuang, pulang, pergi, menangis, bensinnya, ...
3                [isi, pertamax, penuh, ribu, ribu, tau]
4    [andai, gaji, naiknya, serajin, naiknya, harga...
```

Figure 6. Stopword removal result

The last stage of text preprocessing is stemming, which is the process of changing words into their basic forms. The results of the stemming process can be seen in Figure 7.

3.2. Classification Model

The dataset that has been converted into numerical form will then be used to create a classification model; in this stage, training data is used. The classification model aims to predict the labels from the testing data based on the model that has been formed from the training data. In the classification model, the SVM algorithm is used, and the RBF kernel trick is applied to map the dataset into a higher dimensional space.

```
[ ]  diresmikan : resmi
     dipermainkan : main
     licin : licin
     belut : belut
     pemaaf : maaf
     miskinkan : miskin
     judulnya : judul
     magic : magic
     carpet : carpet
     ride : ride
     prediksi : prediksi
     hitungnya : hitung
     komentarnya : komentar
     pemikirannya : pikir
     kecukupan : cukup
     mempertanggung : tanggung
     jawabkan : jawab
     nyimak : nyimak
     gajimu : gaji
     dealer : dealer
     tengok : tengok
     disuguhi : suguh
     terserahlah : serah
     becak : becak
     belasan : bas
     gorengan : goreng
     tanduk : tanduk
```

Figure 7. Stemming result

We chose the RBF kernel because it is particularly well-suited for non-linearly separable data, such as text datasets. With proper tuning of hyperparameters like gamma (γ) and the penalty parameter C, the RBF kernel balances flexibility and regularization, reducing the risk of overfitting. In addition, parameter selection (hyperparameter tuning) is used to improve the performance of the model [25]. We have the pseudocode in this step in the following:

```
x_train, x_test, y_train, y_test ← train_test_split(dataset['tweet_clean'], dataset['label'],test_size= "20%")
vectorizer ← TfidfVectorizer()
x_train ← vectorizer.fit_transform(x_train)
x_test ← vectorizer.transform(x_test)
```

The pseudocode demonstrates the two primary processes: dataset splitting, which divides the text and labels into training and testing data with a ratio of 80:20, and feature extraction, which uses the TF-IDF approach to convert the text into a numerical representation that the machine can utilize.

3.3. Hyperparameter Tuning

At this step, the selection of value pairs for each parameter will be implemented, which can produce the best accurate value. In the RBF kernel two parameters are determined, namely the value of C (Cost) and the value of γ (gamma). Cost C and γ values are obtained from the hyperparameter tuning process manually, where the value of each parameter is inserted one by one from a certain interval until the best model performance is discovered [26]. Table of comparison of the values of each parameter C, γ, accuracy and training time can be seen in the Table 2.

Based on the experiments conducted for each of these parameter values, the parameters utilized were C and γ, with values of 1.1, respectively. These values were chosen due to their ability to yield the highest accuracy, reaching 81%, or equivalently, a value of 0.808 from the model. Following the model's establishment, testing data will be introduced into the model for label prediction. The accuracy value is derived from 2161 iterations of the training model. As the number of iterations increases, accuracy tends to converge at 81%. A graphical representation of the relationship between iterations and accuracy is presented in Figure 8.

Table 2. Comparison of the values of C, γ, accuracy and training time

No	C Value	γ Value	Accuracy	Training Time(s)
1	1	1	0.799	0.321
2	1	1.1	0.803	0.365
3	1	1.2	0.801	0.809
4	1.1	1	0.803	0.303
5	1.1	1.1	0.808	0.547
6	1.1	1.2	0.801	0.337
7	1.2	1	0.796	0.995
8	1.2	1.1	0.796	0.377
9	1.2	1.2	0.796	0.665

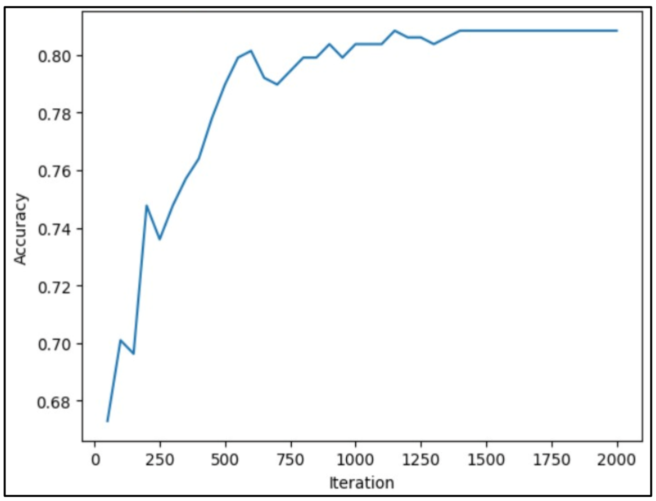

Figure 8. The relationship between iteration and accuracy

3.4. Data Validation

The prediction results obtained from the classification model need to undergo validation to assess the performance and accuracy of the model in predicting labels for testing data. In this stage, k-Fold cross-validation is employed, whereby the training data is partitioned into k segments, each containing an equal number of data samples. Within each segment, different data samples are employed for validation against the testing data. The value of k utilized in this study is set to 5, resulting in five distinct scenarios with varying data samples for each segment. The accuracy achieved in each segment and its overall average are depicted in Figure 9.

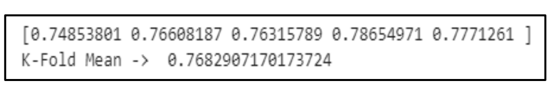

Figure 9. K-fold cross validation result

3.5. Model Testing

At this stage, the classification model will be evaluated using the confusion matrix method. The confusion matrix will produce true positive, true negative, false positive, and false negative values by comparing the actual values and predicted values of the two classes. The comparison between actual values and predicted values is illustrated in Figure 10.

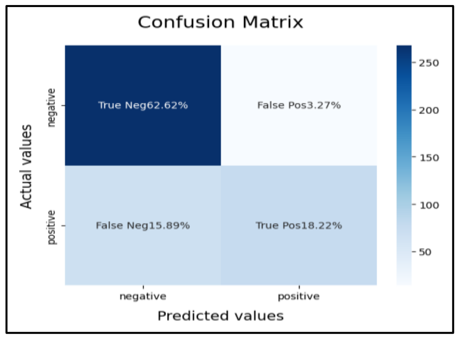

Figure 10. Confusion matrix

3.6. Data Visualization

The visualization results of the classification of public sentiment towards government policy in increasing fuel prices will be displayed in two types of visualization. There is *a word cloud*, which will show the most frequently used words from each class, and *a pie chart,* which will show the percentage of each class. *Word clouds* of the negative class and positive class can be seen in Figures 11 and 12.

Figure 11. Negative Word Cloud

Figure 12. Positive Word Cloud

Figure 13 shows a *pie chart* visualization showing the percentage of each class of comments.

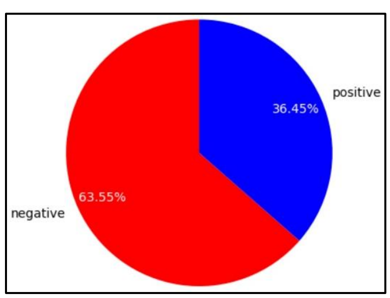

Figure 13. Pie Chart of Positive and Negative Public Responses

3.7. User Interface

The classification model's interface was developed as a straightforward and interactive web-based application. It enables users to input new comment data and subsequently predicts the class of the new data. The outcomes of classifying new comments for each category are depicted in Figures 14 and 15.

Figure 14. Positive class

Figure 15. Negative class

4. DISCUSSION

This study investigated public sentiment on Twitter regarding the government's policy to increase fuel prices by analyzing tweets from verified news accounts. As shown in Table 1, a comprehensive dataset of 5,076 tweets was collected and meticulously filtered to 2,161 relevant entries. The data underwent extensive preprocessing—including case folding, cleansing, tokenization, stopword removal, and stemming—to convert raw textual content into a clean and analyzable format. The transformation of this text into numerical features using TF-IDF vectorization was essential for training the subsequent classification model.

An SVM classifier with an RBF kernel was employed to predict sentiment labels, with careful hyperparameter tuning playing a pivotal role in achieving optimal performance. By manually adjusting the cost (C) and gamma (g) parameters, both of which were determined to be 1.1, the model achieved a robust accuracy of 81% over 2,161 iterations, as shown in Figure 8. The iterative tuning process, combined with K-Fold cross-validation, confirmed the model's stability and reliability, as evidenced by consistent performance across different validation segments. The confusion matrix further provided insights into the distribution of true positives, true negatives, false positives, and false negatives, underscoring the classifier's predictive precision.

Complementing the quantitative results, data visualization through word clouds and pie charts effectively illustrated the frequency and proportion of sentiments expressed in the tweets. The development of an interactive, web-based user interface further enhanced the practical application of this research by enabling real-time sentiment analysis of new comment data. Overall, the findings suggest that with appropriate preprocessing, feature extraction, and parameter optimization, SVM-based classification models can reliably capture and predict public sentiment on social media platforms, thereby offering valuable insights into public opinion on governmental policies.

5. CONCLUSION

With an 80:20 training-testing split, the SVM model employing the RBF kernel achieved an 81% accuracy rate, and an average accuracy of 77% under K-Fold cross-validation. Analysis of 2,161 public comments regarding fuel price hikes in 2022 revealed that 63.55% expressed negative sentiment, while 36.45% were positive. These results indicate a predominance of public disapproval, suggesting that the government's policy may not have been implemented at an optimal time. The SVM algorithm successfully developed a reliable model for classifying public sentiment, offering a data-driven method to assess public reaction to policy changes. These findings highlight the importance of considering public sentiment prior to decision-making, as sentiment analysis can provide valuable insights into societal acceptance. Furthermore, the success of the model underscores the potential for broader applications in various fields, including electoral analysis, market research, and integration into Decision Support Systems (DSS) for enhancing timely and informed policymaking.

ACKNOWLEDGMENT

We express our sincere gratitude to DIPA Universitas Tadulako for their generous support under contract number 1221z/UN28.2/KU/2023. This funding has been instrumental in advancing our research project and making a significant contribution to the growth of our field.

REFERENCES

[1] R. E. Husna, R. Wasono, and M. A. Haris, "Analisis Sentimen pada Twitter Mengenai Netflix Diblokir Telkom Menggunakan Support Vector Machine," in *Prosiding Seminar Nasional VARIANSI*, 2020.

[2] S. Koda, A. Zeggada, F. Melgani, and R. Nishii, "Spatial and Structured SVM for Multilabel Image Classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 10, pp. 5948–5960, 2018, doi: 10.1109/TGRS.2018.2828862.

[3] A. Mishra, ; Sushree, S. Dash, and ; Swati Tiwari, "Unlocking the Magic of Facial Recognition: Empowering Security and Emotions With Svm," *J. Data Acquis. Process.*, vol. 38, no. 2, p. 2402, 2023, doi: 10.5281/zenodo.776955.

[4] Y. Liang *et al.*, "Study on the Prediction of Low-Index Coal and Gas Outburst Based on PSO-SVM," *Energies*, vol. 16, no. 16, 2023, doi: 10.3390/en16165990.

[5] Y. Lin, "Research on HOG-SVM pedestrian detection method based on FPGA," *Appl. Comput. Eng.*, vol. 9, no. 1, pp. 272–281, 2023, doi: 10.54254/2755-2721/9/20230110.

[6] U. Kurniasih and A. T. Suseno, "Analisis Sentimen Terhadap Bantuan Subsidi Upah (BSU) pada Kenaikan Harga Bahan Bakar Minyak (BBM)," *J. Media Inform. Budidarma*, vol. 6, no. 4, p. 2335, 2022, doi: 10.30865/mib.v6i4.4958.

[7] I. Nurhidayat and B. Pimpunchat, "A Comparative Approach to SVM Kernel Functions via Accurate Evaluating Algorithms," *J. Eng. Sci. Technol.*, vol. 18, no. 4, pp. 2078–2090, 2023.

[8] R. Pandey, S. K. Khatri, N. K. Singh, and P. Verma, Eds., *Artificial intelligence and machine learning for EDGE computing*. Academic Press, 2022.

[9] M. Sinambela, M. Situmorang, K. Tarigan, S. Humaidi, and M. Sirait, "Waveforms Classification of Northern Sumatera Earthquakes for New Mini Region Stations Using Support Vector Machine," *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 11, no. 2, pp. 489–494, 2021, doi: 10.18517/ijaseit.11.2.12503.

[10] C. Yan, "Stock Price Prediction of Walmart Based on Combination of SVM and LS-SVM Models," *BCP Bus. Manag.*, vol. 38, pp. 363–371, 2023, doi: 10.54691/bcpbm.v38i.3716.

[11] L. Shao and Y. Gao, "A Gas Prominence Prediction Model Based on Entropy-Weighted Gray Correlation and MCMC-ISSA-SVM," *Processes*, vol. 11, no. 7, 2023, doi: 10.3390/pr11072098.

[12] J. Serin, K. T. Vidhya, I. S. Mary Ivy Deepa, V. Ebenezer, and A. Jenefa, "Gender Classification from Fingerprint Using Hybrid CNN-SVM," *J. Artif. Intell. Technol.*, vol. 4, no. 1, pp. 82–87, 2024, doi: 10.37965/jait.2023.0192.

[13] S. Acharya, T. Kar, U. C. Samal, and P. K. Patra, "Performance Comparison between SVM and LS-SVM for Rice Leaf Disease detection," *EAI Endorsed Trans. Scalable Inf. Syst.*, vol. 10, no. 6, pp. 1–7, 2023, doi: 10.4108/eetsis.3940.

[14] A. S. Abdelfattah, "SVM transformations for Multi-labeled Topics," *Proc. ACM RACS*, vol. 1, no. December 2022, 2022, doi: 10.36227/techrxiv.21671915.

[15] V. Kaushik and M. Kumar, "Water surface profile prediction in non-prismatic compound channel using support vector machine (SVM)," *AI Civ. Eng.*, vol. 2, no. 1, pp. 1–12, 2023, doi: 10.1007/s43503-023-00015-1.

[16] T. Fan, H. Wang, P. Wu, C. Ling, and M. T. Ahvanooey, "Multimodal sentiment analysis for social media contents during public emergencies," *J. Data Inf. Sci.*, pp. 1–27, 2023, doi: 10.2478/jdis-2023-0012.

[17] N. Hendrastuty, A. Rahman Isnain, and A. Yanti Rahmadhani, "Analisis Sentimen Masyarakat Terhadap Program Kartu Prakerja Pada Twitter Dengan Metode Support Vector Machine," *J. Inform. J. Pengemb. IT*, vol. 6, no. 3, pp. 150–155, 2021.

[18] R. Wati and S. Ernawati, "Analisis Sentimen Persepsi Publik Mengenai PPKM Pada Twitter Berbasis SVM Menggunakan Python," *J. Tek. Inform. UNIKA St. Thomas*, vol. 06, pp. 240–247, 2021, doi: 10.54367/jtiust.v6i2.1465.

[19] H. Hu, K. Li, W. Liang, Q. Li, and Z. Xie, "Kilometer Sign Positioning-Aided INS/Odometer Integration for Land Vehicle Autonomous Navigation," *IEEE Sens. J.*, vol. 23, no. 4, pp. 4143–4158, 2023, doi: 10.1109/JSEN.2023.3236063.

[20] D. Darwis, E. S. Pratiwi, and A. F. O. Pasaribu, "Penerapan Algoritma Svm Untuk Analisis Sentimen Pada Data Twitter Komisi Pemberantasan Korupsi Republik Indonesia," *Edutic - Sci. J. Informatics Educ.*, vol. 7, no. 1, pp. 1–11, 2020, doi: 10.21107/edutic.v7i1.8779.

[21] F. S. Pamungkas and I. Kharisudin, "Analisis Sentimen dengan SVM, NAIVE BAYES dan KNN untuk Studi Tanggapan Masyarakat Indonesia Terhadap Pandemi Covid-19 pada Media Sosial Twitter," *Pros. Semin. Nas. Mat.*, vol. 4, pp. 1–7, 2021, [Online]. Available: https://journal.unnes.ac.id/sju/prisma/article/view/45038

[22] F. S. Jumeilah, "Penerapan Support Vector Machine (SVM) untuk Pengkategorian Penelitian," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 1, no. 1, pp. 19–25, 2017, doi: 10.29207/resti.v1i1.11.

[23] A. M. Pravina, I. Cholissodin, and P. P. Adikara, "Analisis Sentimen Tentang Opini Maskapai Penerbangan pada Dokumen Twitter Menggunakan Algoritme Support Vector Machine (SVM)," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 3, no. 3, pp. 2789–2797, 2019, doi: 10.1061/9780784415313.ch06.

[24] L. Q. Zalyhaty, "Analisis Sentimen Tanggapan Masyarakat Terhadap Vaksin Covid-19 menggunakan Algoritma Support Vector Machine (SVM)," Universitas Dinamika, 2021.

[25] A. Julian and R. Devipriya, "Exploring Hyperparameter Tuning Strategies for Optimizing Model Performance," in *2024 International Conference on Trends in Quantum Computing and Emerging Business Technologies*, 2024, pp. 1–4. doi: 10.1109/TQCEBT59414.2024.10545054.

[26] J. A Ilemobayo *et al.*, "Hyperparameter Tuning in Machine Learning: A Comprehensive Review," *J. Eng. Res. Reports*, vol. 26, no. 6, pp. 388–395, 2024, doi: 10.9734/jerr/2024/v26i61188.