IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 13, No. 4, December 2024, pp. 4948~4956
ISSN: 2252-8938, DOI: 10.11591/ijai.v13.i4.pp4948-4956

Indonesian sentiment towards global economic recession in 2023 using optimized hyperparameters of support vector machine kernels

Dairatul Maarif, Adinda Aulia Hafizha, Andi Kurniawan
Department of International Relations, Faculty of Social and Political Sciences, Universitas Pembangunan Nasional Veteran Jakarta, Jakarta, Indonesia

Article Info

Article history:

Received Jun 14, 2023
Revised Mar 6, 2024
Accepted Apr 14, 2024

Keywords:

Global recession
Kernels
Optimized hyperparameters
Sentiment analysis
Support vector machine

ABSTRACT

The potential for the 2023 global recession has troubled people worldwide, particularly in light of the COVID-19 pandemic. This study employs a sentiment analysis approach to examine how the Indonesian internet community, particularly on Twitter, perceives the topics related to the global economic recession. We collected 11,017 uploaded tweets that were analyzed using support vector machine classifier with linear, radial basis function (RBF), sigmoid, and polynomial kernel schemes. Furthermore, we optimized the classifiers with C, Gamma, and degree hyperparameters. Empirical evidence indicates a lack of preparedness to face a global recession, evidenced by most responses towards 2023 global recession exhibiting concerns about high inflation and economic instability. The finding also suggests that the optimized RBF is a superior modeling kernel relative to others. Collectively, these results provide insights with significant implications for sentiment analysis, natural language processing, and the study of behavioural economics.

Corresponding Author:

Adinda Aulia Hafizha
Department of International Relations, Faculty of Social and Political Sciences
Universitas Pembangunan Nasional Veteran Jakarta
RS. Fatmawati Street, Pondok Labu, South Jakarta 12450, Indonesia
Email: hafizhadinda@gmail.com

1. INTRODUCTION

In recent years, economic recession has had significant negative impacts on the global community, notably highlighted by the far-reaching consequences of the 2008 global financial crisis on economies worldwide. Since 2019, the COVID-19 pandemic has directly impacted various economic aspects, given that the related shutdowns mandated by the government produced a momentous shock to the global economy as well as an excessive loss of employment that heavily affected the working class [1]. Furthermore, given the tightening financial conditions in many regions, debates regarding global economic recession are inevitable, especially in Indonesia with 277 million inhabitants.

Debates regarding the 2023 global economic recession in Indonesia are fiercely argued in Twitter. Some Twitter users feel optimistic that Indonesia will not be subjected to economic recession in 2023 because few economic indicators for Indonesia show promising results. However, according to others, Indonesia will be subject to progressively “dark” global uncertainties in 2023, conveyed by the decrease in people’s purchasing power and the Indonesia composite index (ICI) due to the COVID-19 pandemic. This controversial debate continues to spread, which in turn formed various public opinions with conflicting views. Economists and decision-makers may find this debate concerning since public opinion has the power to impact macroeconomics and the state of the world economy. This supports Akerlof's contention that the inclusion of social and psychological viewpoints would facilitate a better understanding of some economic events [2].

Previous studies have examined public opinions on economic issues. Wielen and Barrios [3] used data from Google Trends to identify a significant change in economic attitudes during the COVID-19 crisis and subsequent lockdown in the European Union. Additionally, Carosia *et al.* [4] analyzed sentiments related to Brazilian stock market movements on Twitter using various machine learning techniques. The study suggests a correlation between the sentiment expressed on social media and the fluctuations in stock prices in the financial market. Swathi *et al*. [5] also predicted stock prices using Twitter sentiment analysis. Carosia *et al.* [4] found that multilayer perceptrons (MLP) were the best model for predicting stock prices. However, Swathi *et al*. [5] has shown that a novel teaching and learning based optimization (TLBO) model with long short-term memory (LSTM) achieved even greater accuracy, reaching 94.73%. Cruz *et al*. [6] conducted research implementing lexicon-based analysis. The study presents substantial evidence supporting the assumption that Twitter messages altered financial indices during the COVID-19 and H1N1 pandemics, with notable effects in the case of COVID-19.

Several studies have utilized sentiment analysis techniques, including deep learning, MLP, and lexicon-based methods, to capture public opinions on economic issues [3]–[6]. However, no similar study has been conducted to analyze public opinions on the 2023 economic recession, particularly among Indonesians. Therefore, this research aims to identify how the Indonesian internet community, particularly on Twitter, perceives the topics related to the global economic recession. Additionally, this research will utilize support vector machine (SVM) kernels that have been improved through hyperparameters optimization. The chosen method was selected for their potential to enhance the performance of the SVM classification model, either by utilizing kernels [7]–[13] or implementing hyperparameters optimization [14], [15]. Therefore, to increase the novelty and value of this research, we combined kernels and hyperparameters optimization to improve the accuracy of the results. The study’s findings will have practical implications for economic policymakers to tailor communication strategies, economic policies, and crisis management approaches to reflect public sentiment during the global economic recession. This study also enhances the methodological aspects of sentiment analysis by showcasing the significance of optimizing SVM hyperparameters for contextual domains.

2. METHOD

2.1. Dataset

This research uses Indonesians tweets from Twitter as a dataset, ranging from 1 January 2022 to 31 December 2022. Figure 1 outlines the steps taken in this study based on the general framework used to analyze textual data with the addition of hyperparameters optimization step [14]. This work is conducted using Python programming language, version 3.7.0. Using the keywords ‘Indonesia global recession 2023’, we successfully crawled 11.017 tweets with the Tweepy library. After crawling from Twitter, the fresh-mined dataset was further processed with text preprocessing step, which consisted of text cleaning with regular expressions (RE), case folding, tokenization with nltk.tokenize function, stopword removal, and stemming with Porter stemmer [16].

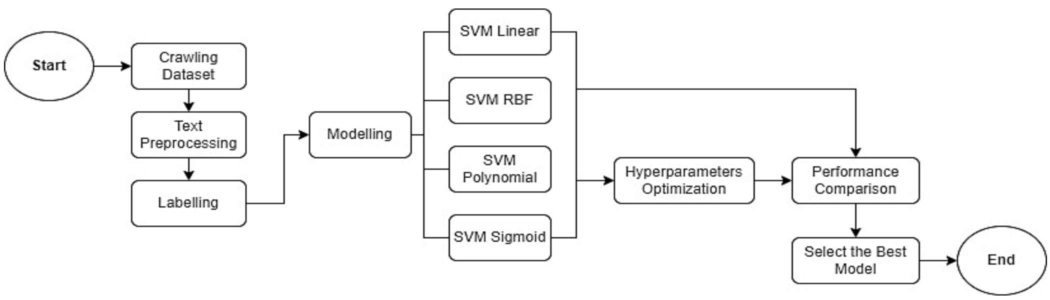

Figure 1. The flowchart of sentiment analysis research using SVM kernels with hyperparameters optimization

After the preprocessing step, the cleaned dataset was divided into train data and test data with a percentage ratio of 80:20. Then, the train data was labeled into two features, which were positive and negative, by utilizing TextBlob. This method ensures more accurate and precise results in sentiment analysis cases, outperforming AFINN and valence aware dictionary for sentiment reasoning (VADER) [17]. When the trained data has been annotated with TextBlob, the next step is to implement the vectorization stage with term frequency-inverse document frequency (TF-IDF). TF-IDF is recognized as the most robust feature extraction ever used for clustering and classifying textual data [18]–[20], outperforming bag of words [21], and word2vec [22]. We utilize the TfidfVectorizer library with parameters max_features (2,000), min_df (5), max_df (0.7), stopwords.words ('English'), and ngram_range (1.3). Following the vectorization step, the research process will continue with model construction.

2.2. Model construction and hyperparameters optimization

After preprocessing, labeling, and vectorization, classification models were built using SVM kernel functions. Typically, kernel functions project data from its original space into a higher-dimensional space, making it possible to separate the data using a linear decision boundary [23], [24]. This research utilized the scikit-learn library by calling "from sklearn.svm import SVC" command to assign each kernels into our model, as seen in Table 1 [25]. Afterward, we add kernel as parameters in the SVC function, with its default setting.

Table 1. Types of kernels in SVM

Kernels	Functions [26]	
Linear	$K(x_i, x_j) = \langle x_i, x_j \rangle$	(1)
Radial basis function (RBF)	$K(x_i, x_j) = e^{-\gamma \lVert x_i - x_j \rVert^2}, \gamma > 0$	(2)
Polynomial	$K(x_i, x_j) = (\langle x_i, x_j \rangle + r)^d, d > 0, r \geq 0$	(3)
Sigmoid	$K(x_i, x_j) = tanh(\gamma(x_i - x_j) + r)$	(4)

However, SVM kernel functions have limitations that should be considered [23]. Firstly, several studies have implied that although the datasets used in various research of SVM kernels are consistent (tweets data), the optimization outcomes of the utilized model yield distinct results. Additionally, it lacks efficiency in handling multiclass problems. To address the issue of inefficiency in multiclass tasks, this study utilized only two sets of labels: positive and negative sentiments. In addition to using kernels, this research also applies hyperparameters optimization. Hyperparameters are variables that affect the learning process and influence the model parameters learned by an algorithm [14]. Table 2 shows the hyperparameters of each SVM kernel applicable for optimization. Then, using grid search, hyperparameters were optimized by preparing the machine learning algorithm for every conceivable combination of parameters [14], [15].

Table 2. Hyperparameters optimization params in SVM

Kernels	C Score (Regularisation)	Gamma	Degree
Linear	[0.001, 0.01, 0.1, 1, 10, 100, 1000]	-	-
RBF	[0.001, 0.01, 0.1, 1, 10, 100, 1000]	[1, 0.1, 0.01, 0.001, 0.0001, scale, auto]	-
Polynomial	[0.001, 0.01, 0.1, 1, 10, 100, 1000]	[1, 0.1, 0.01, 0.001, 0.0001, scale, auto]	[2, 3, 4, 5]
Sigmoid	[0.001, 0.01, 0.1, 1, 10, 100, 1000]	[1, 0.1, 0.01, 0.001, 0.0001, scale, auto]	-

Nevertheless, it should be noted that hyperparameters are not prone to overfitting issues [27]. To handle this problem, this research uses cross validation (CV). CV can identify a stable optimum that performs well in most subsets, rather than a sharp optimum that only performs well in a single validation set. To mitigate this, we repeat each iteration process of finding the best parameters in the predefined SVM kernels. Here, we utilize GridSearchCV function from sklearn.model_selection and defined the parameters range of each kernels (C, Gamma, Degree) in param_grid variable. Then, we run GridSearchCV with refit=True and verbose=3. After building the model, the next step is validating the models that have been built to find the best performance of these models with confusion matrix. In the confusion matrix, the classification process outcomes are denoted by four terms: true positive (TP), true negative (TN), false positive (FP), and false negative (FN). Table 3 illustrates the performance parameters in the confusion matrix.

Table 3. Confusion matrix

Predicted/Actual classes	Predicted class (positive)	Predicted class (negative)
Actual class (positive)	TP	FN
Actual class (negative)	FP	TN

Each performance metric in the confusion matrix is further calculated to determine the accuracy of the model, using indicators such as accuracy, precision, recall, and F1 score, which equations are shown in (5) to (8):

$$\text{Accuracy}=\frac{TP + TN}{TP + TN + FP + FN}\ x\ 100\% \tag{5}$$

$$\text{Recall}=\frac{TP}{FN + TP}\ x\ 100\% \tag{6}$$

$$\text{Precision}=\frac{TP}{FP+TP}\ x\ 100\% \tag{7}$$

$$\text{F1 Score}=2x\,\frac{Precision\ x\ Recall}{Precision+Recall} \tag{8}$$

3. RESULTS AND DISCUSSION

3.1. Results of sentiment classification

The study's initial objective was to identify the sentiments regarding the 2023 economic recession in Indonesia through Twitter. We utilize R software's GGplot package for data visualization to enhance comprehension of the obtained results. Using the time-series function, we map the timeline of Indonesian sentiments, encompassing the highest and the lowest perceptions of each sentiment throughout 2022. This step resulted in Figure 2, which presents Indonesians perceptions of global recession through comments on Twitter. To aid in presenting the results and observing trends, we categorized the sentiment analysis by quarter, specifically Q1 (January-March), Q2 (April-June), Q3 (July-September), and Q4 (October-December).

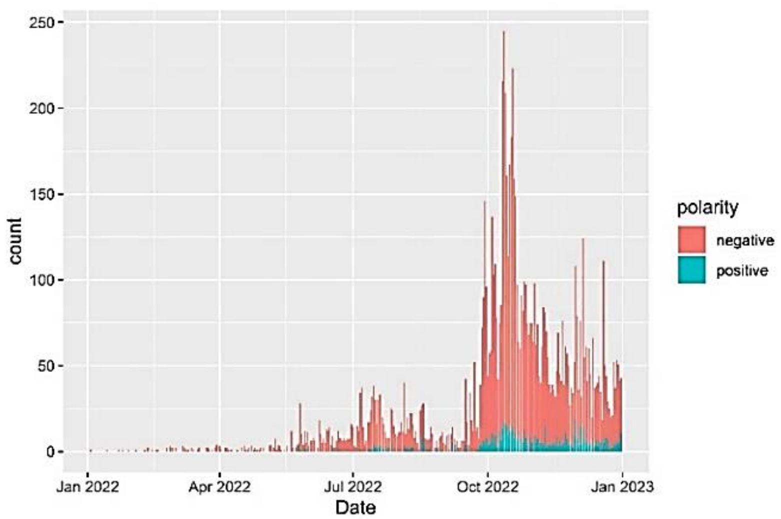

Figure 2. Dynamics of comments expressing different sentiments towards global recession in 2023

Sentiments of the Indonesians regarding the global recession are fluctuating. During the first quarter (Q1) of 2022, both positive and negative sentiments were relatively insignificant and seldom discussed by Twitter users in the country, as shown by the negative bars that did not exceed 50 in counts. Meanwhile, sentiments increased during the last month of the second quarter (Q2) of 2022. An explanation for this might be that the possibility of a global recession began to be discussed in various online media. Meanwhile, at the beginning of Q3, negative sentiments suddenly increased compared to the total of sentiments at the end of Q2. However, it quickly dissipated in September, as shown by the decrease in sentiment graph during Q3. The sentiments peak was during October 2022, when the negative sentiments reached 200-250 tweets per day. It has also observed a significant increase in positive sentiments, reaching more than 20 tweets during October. This finding may indicate that Indonesians are increasingly aware of the impending economic recession in 2023. The revelation has caused widespread panic, prompting debates on whether Indonesia's economy is susceptible to recessionary pressures.

We can turn to a deeper analysis of public sentiment towards recession in 2023 by examining negative and positive tweets using the most frequently used words. Figure 3 highlights the five most-used words in tweets classified as negative sentiment. From the Figure 3, it is apparent that Indonesian society's negative sentiment towards recession is circulating the financial impact of it. A further inference that can be drawn from the information presented is that Indonesians view economic recession as a menacing force. This unfavorable outlook stems from anxiety over the potential for job loss and ongoing destitution, which unfortunately remains unsteady in some recession-impacted nations. According to Ssenyonga and Shafiullah [28], the pandemic has led to a decline in output across most Indonesian economic sectors, resulting in a surge in unemployment from 5.28% in the third quarter of 2019 to 7.07% in the third quarter of 2020. Consequently, the increasing levels of unemployment and bankruptcy caused by the COVID-19 pandemic, combined with the global recession, have the potential to destabilize the country. Therefore, given the widespread recognition that the global economic crisis worsens poverty, creates inequality, and decreases life expectancy, it is unsurprising that a recession can elicit unfavorable reactions within the population.

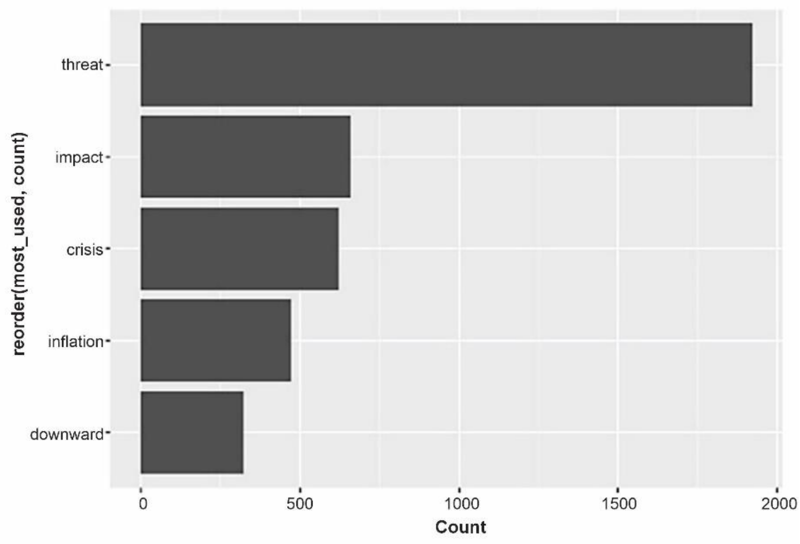

Figure 3. Most common words used in negative sentiments

Having discussed the negative sentiments on global recession, the next section in this paper will discuss the positive sentiments. Figure 4 illustrates public opinion toward a recession that is classified as positive. Based on the bar chart, we can conclude that "work," "potential," and "growth" are the most used words in the tweets of Indonesian society regarding the debate on the global economic recession. A minority within Indonesian society posits that the country will not suffer significantly from the global recession. They contend that through continued work on potential resources, the nation may withstand these looming hazards and even see a positive impact on gross domestic product (GDP). By following these measures, the optimist society feels more confident in Indonesia's ability to withstand a potential economic downturn. Other than utilizing the potential resources, it is worth noting that a small part of Indonesian society also used 'state-owned enterprises' in correlation to the positive sentiments. This interesting result might be explained by the fact that state-owned enterprises serve as a model for the nation's economy since various sectors of state-owned enterprises activities contribute to achieving the country's lofty goals. This is in line with the objective of Indonesian state-owned enterprises, which is to provide public services to aid the government in promoting the well-being and prosperity of the population [29]. Given the importance of state-owned enterprises in the Indonesian economy, it is not unexpected that people believe the local economy may survive and grow exponentially through the industrial advancement of state-owned enterprises.

3.2. Results of model construction

Table 4 displays the values of model validation using four kernels of SVM without optimized hyperparameters. From this finding, the best model for the Twitter dataset is linear kernel, with an accuracy score of 93.73%. This finding is consistent with [7], [8]. Both researchers stated that the linear kernel of SVM with default parameters can generate results above 90%, making it highly appropriate for the Twitter dataset. A possible explanation for this supportive result is the dataset itself, where the emotions contained in Twitter data can be separated linearly due to the presence of only two features: positive and negative.

Therefore, from this result, the linear kernel without optimized hyperparameters are suitable for textual datasets with emotional features.

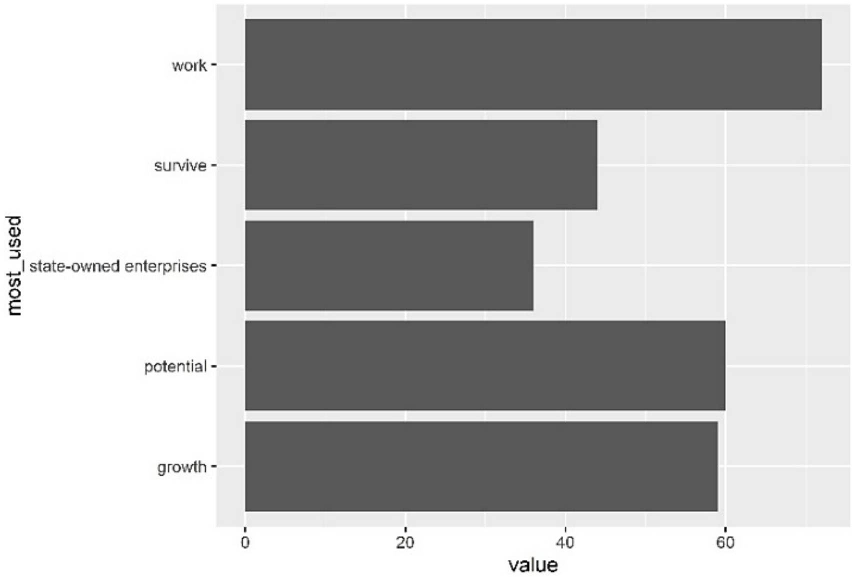

Figure 4. Most common words in positive sentiments

Table 4. Model validation values using four kernels of SVM

Kernels	Accuracy (%)	Precision (%)	Recall (%)	F1 score (%)
Linear	93.73	13.01	84.21	22.54
RBF	93.55	8.13	100.00	15.04
Polynomial	93.67	10.57	92.86	18.98
Sigmoid	93.61	9.76	92.31	17.65

Meanwhile, Table 5 presents the model validation values of each kernel in SVM after hyperparameters optimization. Surprisingly, a closer inspection of the table shows that the optimized RBF is the best model for the previously labeled tweets dataset, with a higher accuracy score than the default one, which reached 94.35%. Possible explanation for this result is that the improved RBF (optimized RBF) works well in high-dimensional space data. It is also apparent that the optimized RBF is the most suitable algorithm for the research because of it is capability to handle immense datasets, matching the characteristics of k-nearest neighbors (k-NN) algorithm, which is considered efficient against large amounts of data, and known for its ability to withstand noise. This finding reflect those of [23] who also found that when the dataset is large, the RBF kernel function performs more accurately than the polynomial kernel function for classification tasks. Although there is no standardized definition of what constitutes 'large data', we assume that the data in this research is considered large because it exceeds 10,000 records.

Table 5. Model validation values after hyperparameters optimization using four kernels of SVM

Kernels	Accuracy (%)	Precision (%)	Recall (%)	F1 score (%)
Linear	93.73	13.01	84.21	22.54
RBF	94.35	24.39	83.33	37.74
Polynomial	93.73	11.38	93.33	20.29
Sigmoid	93.73	13.01	84.21	22.54

This outcome varies from the default parameter model. This discrepancy could be attributed to the effect of hyperparameters optimization, which is proven effective in optimizing the default model. If we take a closer look at the comparison of default and post-tuned models of SVM kernels, almost all accuracy, precision, recall, and F1 score show a rather significant increase or remain stagnant after hyperparameters optimization were implemented. The increase in each model validation has also proven that implementing optimized hyperparameters will improve the model performance or, at the very least, remain stagnant in both post-tuned and pre-tuned models. The increased accuracy score in the SVM kernels in this study corroborates the earlier findings [14], [15], where it has been stated that the optimized SVM with C, Gamma, and degree parameters worked best towards the tweets dataset. However, we can observe a slight decrease on recall scores in the tuned SVM kernel models. This could have happened because of the imbalanced dataset, where negative sentiments are higher than positive ones, so the model may exhibit bias toward predicting the majority class, leading to a lower recall for the minority class.

Considering another measurement of supervised machine learning model performance, Figure 5 displays the confusion matrix of the improved RBF kernel algorithms. Upon closer inspection of the matrix, the optimized SVM-RBF model exhibits moderate errors in quadrants two and three, but higher values in quadrant four, indicating TP. Therefore, we can conclude that the optimized SVM-RBF model still has relatively minor errors. The reason for these errors may be due to an imbalanced dataset. As a result, the optimized model had low TP and TN values for the separated classes. To address this issue, future research can use a more balanced dataset, incorporate other case studies, set an appropriate threshold for the model, experiment with additional feature extraction techniques, implement dataset classification, and optimize the SVM kernel research with hyperparameters to create a better classifier.

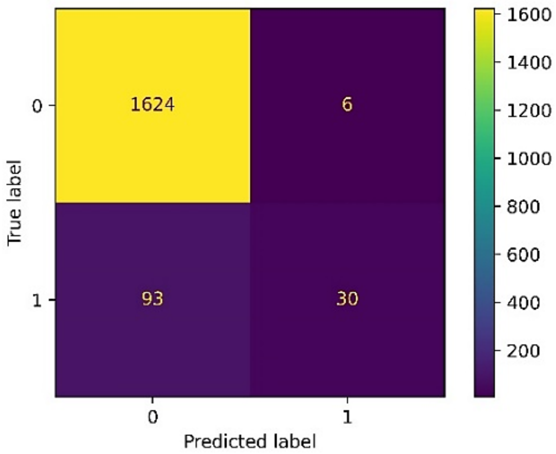

Figure 5. Confusion matrix SVM-RBF kernel’s optimized model

4. CONCLUSION

Based on the conducted research, the results indicate that the Indonesians holds mostly negative viewpoints on the global economic recession in 2023. The most commonly used words were ‘threat,’ ‘crisis,’ and ‘inflation’. These findings suggest that the public needs to be financially prepared to deal with the recession due to concerns about the potential for high inflation, which could lead to prolonged poverty and joblessness. Meanwhile, the optimized SVM-RBF model was found to be the best model. This supports previous studies’ findings that implementing optimized hyperparameters can improve model performance or, at the very least, maintain it in both post-tuned and pre-tuned models. The study demonstrates through empirical evidence that the optimized hyperparameters technique can yield satisfactory results in sentiment analysis research. This establishes a foundation for future research in an alternative method of optimizing SVM kernels. The practical implications of this study include opportunities for improving policymaking towards crises by analyzing public demands, providing an indirect measure of social and economic stability, and understanding consumer behavior amidst economic recession. Although there was a slight decline in Recall values, this may be due to the imbalanced dataset not working well when kernels were optimized. Future research can address this shortcoming by using a more balanced dataset, utilizing other case studies, and setting an appropriate threshold for the model.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the support of Universitas Pembangunan Nasional Veteran Jakarta for this research.

REFERENCES

[1] R. V. Burkhauser, K. Corinth, and D. H. -Eakin, "Policies to help the working class in the aftermath of COVID-19: lessons from the great recession," *The ANNALS of the American Academy of Political and Social Science*, vol. 695, no. 1, pp. 314–330, 2021, doi: 10.1177/00027162211031772.

[2] G. Akerlof, "Behavioral macroeconomics and macroeconomic behavior," *American Economic Review*, vol. 92, no. 3, pp. 411–433, 2002.

[3] W. V. D. Wielen and S. Barrios, "Economic sentiment during the COVID pandemic: Evidence from search behaviour in the EU," *Journal of Economics and Business*, vol. 115, 2021, doi: 10.1016/j.jeconbus.2020.105970.

[4] A. E. O. Carosia, G. P. Coelho, and A. E. A. Silva, "Analyzing the Brazilian financial market through portuguese sentiment analysis in social media," *Applied Artificial Intelligence*, vol. 34, no. 1, pp. 1–19, 2020, doi: 10.1080/08839514.2019.1673037.

[5] T. Swathi, N. Kasiviswanath, and A. A. Rao, "An optimal deep learning-based LSTM for stock price prediction using twitter sentiment analysis," *Applied Intelligence*, vol. 52, no. 12, pp. 13675–13688, 2022, doi: 10.1007/s10489-022-03175-2.

[6] D. V. -Cruz, V. F. -Cortez, A. L. -Chau, and R. S. -Almazán, "Does Twitter affect stock market decisions? financial sentiment analysis during pandemics: a comparative study of the H1N1 and the COVID-19 periods," *Cognitive Computation*, vol. 14, no. 1, pp. 372–387, 2022, doi: 10.1007/s12559-021-09819-8.

[7] D. F. Sengkey, A. Jacobus, and J. M. Fabian, "Effects of kernels and the proportion of training data on the accuracy of SVM sentiment analysis in lecturer evaluation," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 9, no. 4, pp. 734–743, Dec. 2020, doi: 10.11591/ijai.v9.i4.pp734-743.

[8] L. K. Ramasamy, S. Kadry, Y. Nam, and M. N. Meqdad, "Performance analysis of sentiments in Twitter dataset using SVM models," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 11, no. 3, pp. 2275–2284, Jun. 2021, doi: 10.11591/ijece.v11i3.pp2275-2284.

[9] M. Rahardi, A. Aminuddin, F. F. Abdulloh, and R. A. Nugroho, "Sentiment analysis of Covid-19 vaccination using support vector machine in Indonesia," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 6, pp. 534–539, 2022, doi: 10.14569/IJACSA.2022.0130665.

[10] H. Syahputra, "Sentiment analysis of community opinion on online store in indonesia on twitter using support vector machine algorithm (SVM)," *Journal of Physics: Conference Series*, vol. 1819, no. 1, Mar. 2021, doi: 10.1088/1742-6596/1819/1/012030.

[11] R. N. Chory, M. Nasrun, and C. Setianingsih, "Sentiment analysis on user satisfaction level of mobile data services using support vector machine (SVM) algorithm," in *Proceedings-2018 IEEE International Conference on Internet of Things and Intelligence System, IOTAIS 2018*, IEEE, 2019, pp. 194–200. doi: 10.1109/IOTAIS.2018.8600884.

[12] M. Alkaff, A. R. Baskara, and Y. H. Wicaksono, "Sentiment analysis of Indonesian movie trailer on YouTube using delta TF-IDF and SVM," in *2020 5th International Conference on Informatics and Computing, ICIC 2020*, 2020. doi: 10.1109/ICIC50835.2020.9288579.

[13] N. P. Arthamevia, Adiwijaya, and M. D. Purbolaksono, "Aspect-based sentiment analysis in beauty product reviews using TF-IDF and SVM algorithm," in *2021 9th International Conference on Information and Communication Technology (ICoICT)*, 2021, pp. 197–201. doi: 10.1109/ICoICT52021.2021.9527489.

[14] K. S. Chong and N. Shah, "Comparison of naive Bayes and SVM classification in grid-search hyperparameter tuned and non-hyperparameter tuned healthcare stock market sentiment analysis," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 12, pp. 90–94, 2022, doi: 10.14569/IJACSA.2022.0131213.

[15] M. Ahmad, S. Aftab, M. S. Bashir, and N. Hameed, "Sentiment analysis using SVM: A systematic literature review," *International Journal of Advanced Computer Science and Applications*, vol. 9, no. 2, pp. 182–188, 2018, doi: 10.14569/IJACSA.2018.090226.

[16] G. Miner, D. Delen, J. Elder, A. Fast, T. Hill, and R. Nisbet, "Chapter 3-conceptual foundations of text mining and preprocessing steps," in *Practical Text Mining and Statistical Analysis for Non-structured Text Data Applications*, G. Miner, D. Delen, J. Elder, A. Fast, T. Hill, and R. A. B. T.-P. T. M. and S. A. for N. T. D. A. Nisbet, Eds., Boston: Academic Press, 2012, pp. 43–51. doi: 10.1016/B978-0-12-386979-1.00003-7.

[17] W. Aljedaani *et al.*, "Sentiment analysis on Twitter data integrating TextBlob and deep learning models: The case of US airline industry," *Knowledge-Based Systems*, vol. 255, 2022, doi: 10.1016/j.knosys.2022.109780.

[18] M. Ali and Y. Mohamed, "A method for clustering unlabeled BIM objects using entropy and TF-IDF with RDF encoding," *Advanced Engineering Informatics*, vol. 33, pp. 154–163, Aug. 2017, doi: 10.1016/j.aei.2017.06.005.

[19] M. Kamyab, G. Liu, and M. Adjeisah, "Attention-based CNN and Bi-LSTM model based on TF-IDF and GloVe word embedding for sentiment analysis," *Applied Sciences*, vol. 11, no. 23, 2021, doi: 10.3390/app112311255.

[20] R. S. Patil and S. R. Kolhe, "Supervised classifiers with TF-IDF features for sentiment analysis of Marathi tweets," *Social Network Analysis and Mining*, vol. 12, no. 1, Dec. 2022, doi: 10.1007/s13278-022-00877-w.

[21] S. Akuma, T. Lubem, and I. T. Adom, "Comparing bag of words and TF-IDF with different models for hate speech detection from live tweets," *International Journal of Information Technology*, vol. 14, no. 7, pp. 3629–3635, Dec. 2022, doi: 10.1007/s41870-022-01096-4.

[22] D. E. Cahyani and I. Patasik, "Performance comparison of tf-idf and word2vec models for emotion text classification," *Bulletin of Electrical Engineering and Informatics*, vol. 10, no. 5, pp. 2780–2788, 2021, doi: 10.11591/eei.v10i5.3157.

[23] A. Patle and D. S. Chouhan, "SVM kernel functions for classification," in *2013 International Conference on Advances in Technology and Engineering (ICATE)*, IEEE, Jan. 2013, pp. 1–9. doi: 10.1109/ICAdTE.2013.6524743.

[24] I. S. Al-Mejibli, J. K. Alwan, and D. H. Abd, "The effect of gamma value on support vector machine performance with different kernels," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 10, no. 5, pp. 5497–5506, Oct. 2020, doi: 10.11591/ijece.v10i5.pp5497-5506.

[25] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, pp. 273–297, 1995.

[26] G. N. Kouziokas, "SVM kernel based on particle swarm optimized vector and Bayesian optimized SVM in atmospheric particulate matter forecasting," *Applied Soft Computing*, vol. 93, Aug. 2020, doi: 10.1016/j.asoc.2020.106410.

[27] L. Yang and A. Shami, "On hyperparameter optimization of machine learning algorithms: Theory and practice," *Neurocomputing*, vol. 415, pp. 295–316, Nov. 2020, doi: 10.1016/j.neucom.2020.07.061.

[28] M. Ssenyonga and M. Shafiullah, "Imperatives for post COVID-19 recovery of Indonesia's education, labor, and SME sectors," *Cogent Economics and Finance*, vol. 9, no. 1, Jan. 2021, doi: 10.1080/23322039.2021.1911439.

[29] Y. S. Rangkuti, B. Nasution, H. Juwana, and M. Siregar, "Liability of state-owned holding company in facing the impact of COVID-19 pandemic," *Russian Law Journal*, vol. 11, no. 1, pp. 1–8, 2023, doi: 10.52783/rlj.v11i1.286.

BIOGRAPHIES OF AUTHORS

Dairatul Maarif holds a Master of Arts degree from Gajah Mada University in Indonesia in 2013. He is an assistant professor at International Relations Department in Universitas Pembangunan Nasional Veteran Jakarta, South Jakarta, Indonesia. His research interests are in big data in international relations, international political economy, and environmentalism in international relations. He has participated in several national accredited researches including community services and scientific workshop. He can be contacted at email: dairatul@upnvj.ac.id.

Adinda Aulia Hafizha is currently a research assistant to the lecturers in International Relations major of Universitas Pembangunan Nasional Veteran Jakarta. Her research interests are big data in international relations, natural language processing, machine learning, and global political economy. She can be contacted at email: hafizhadinda@gmail.com.

Andi Kurniawan received the Master of Science degree in University of Indonesia in 2014. He is currently lecturing at International Relations Department in Universitas Pembangunan Nasional Veteran Jakarta, South Jakarta, Indonesia. His research areas of interest include political economy of natural resources, big data in international relations, political economy of multinational corporations, and methods of quantitative research. He can be contacted at email: andi.k@upnvj.ac.id.