Journal of Information Systems and Informatics
Vol. 6, No. 2, June 2024 e-ISSN: 2656-4882 p-ISSN: 2656-5935
DOI: 10.51519/journalisi.v6i2.713 Published By DRPM-UBD

Sentiment Analysis of Unemployment in Indonesia During and Post COVID-19 on X (Twitter) Using Naïve Bayes and Support Vector Machine

Putu Ayulia Setiawati[1], I Made Agus Dwi Suarjaya[2], I Nyoman Prayana Trisna[3]

[1,2,3]Information Technology Department, Udayana University, Bali, Indonesia
Email: [1]setiawati.ayulia@gmail.com, [2]agussuarjaya@it.unud.ac.id, [3]prayana.trisna@unud.ac.id

Abstract

The COVID-19 pandemic has impacted health, economy, and society. Social distancing measures and quarantine policies have restricted economic activities, leading to downturns in COVID-19-affected regions and a subsequent rise in unemployment rates, particularly in urban areas. Concurrently, there has been a remarkable surge in the utilization of the X (Twitter) platform, with Indonesia ranking 6th globally in X (Twitter) users. This study aims to understand the diverse perspectives of society on unemployment and the factors influencing society's views on unemployment through sentiment analysis of X (Twitter) data. By analyzing 576,764 tweets from April 2020 to October 2023, tweets are categorized into positive, neutral, and negative classes. Classification model was built to classify tweet data by implementing TF-IDF for word weighting, and a pair of machine learning algorithms, Naïve Bayes and Support Vector Machine (SVM). Model evaluation yielded the highest accuracy of 81.5% using Naïve Bayes. The classification outcomes highlight prevalent negative perceptions of unemployment among Indonesians, totaling 50.03%. This research contributes to the literature by providing a large-scale analysis of social media data to uncover public sentiment trends and offering insights for policymakers to address unemployment and improve welfare.

Keywords: Unemployment, tweet, TF-IDF, Naïve Bayes, Support Vector Machine (SVM)

1. INTRODUCTION

The World Health Organization (WHO) announced that the world is facing a global emergency caused by a disease known as COVID-19. The COVID-19 pandemic and its aftermath have posed significant challenges, particularly concerning employment and income for millions of individuals, social security systems, income support mechanisms, the increased responsibilities borne by women, the difficulties faced by migrants and informal sector workers, mental health concerns, and restrictions on economic activities, including the cessation of production and the inability of firms to market their goods and services.

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

Worldwide, 3.3 billion individuals, representing 81% of the global workforce, were impacted by the lockdown. Among this affected workforce, 61% were employed in the informal sector, with 90% of these informal workers residing in low- and middle-income countries [1].

The first case of COVID-19 in Indonesia was recorded in March 2020 in the city of Depok. Its spread was rapid, with the number of cases surpassing 1,500 within a month. This has resulted in wide-ranging impacts not only on health but also on the economic and social sectors. Social distancing policies and quarantine measures have led to a decrease in economic growth and an increase in unemployment rates, especially in urban areas [2]. Badan Pusat Statistik (BPS) documented a rise in the count of unemployment individuals, escalating from 6.82 million individuals in 2019 to 6.88 million individuals in 2020 [3]. Compared to February 2019, there was an addition of approximately 1.2 million people experiencing unemployment in early 2023.

During the pandemic, the use of platform X (Twitter) has drastically increased. In April 2020, the daily active users reached 166 million, and rose to 199 million in 2021 [4]. X (Twitter) has become a platform for individuals to express their concerns and experiences. In Indonesia, X (Twitter) has 14.75 million users, ranking sixth globally. This platform is considered a significant source of big data used for sentiment analysis, aiding in understanding public perceptions of various events and issues.

Several studies have employed machine learning to conduct sentiment analysis on various issues. Satria's research assessed public concern regarding plastic waste through the Support Vector Machine (SVM) approach, revealing DKI Jakarta as the region exhibiting the highest level of concern regarding this matter [5]. Another study by Elisabet found negative perceptions towards regional elections (PILKADA) with Naïve Bayes outperforming SVM [6]. Conversely, in a study in 2023, neutral sentiments were found to dominate in online learning, with SVM outperforming Naive Bayes [7]. In a study by Rizal and Sulastri, the dominance of neutral opinions in online learning was found, with Naive Bayes outperforming SVM [8]. A study on the unemployment rate in India using Lexicon technique found negative sentiments towards jobs in India despite the stable unemployment rate [9].

Most existing studies have focused on different topics and there is limited research specifically addressing how the Indonesian public perceives unemployment during and after the pandemic. This study aims to fill this gap by employing Naïve Bayes and Support Vector Machine (SVM) methods to classify the sentiment of Indonesian society regarding unemployment during and after COVID-19. By comparing these methods, the research seeks to determine which is more effective

in sentiment classification and understand the diverse perspectives of society on unemployment and the factors influencing society's views on unemployment.

Naïve Bayes was selected for its robust capability to classify data effectively, even with a limited amount of training data. Conversely, SVM is known for its high generalization capabilities and ability to classify data well even when trained with a smaller dataset. Although SVM's training time is generally slower, its proficiency in handling non-linear and complex data results in higher accuracy [7]. Prior studies have demonstrated the effectiveness of Naïve Bayes and SVM individually for various sentiment analysis tasks, but they have not specifically addressed unemployment during the COVID-19 pandemic or compared the performance of these algorithms in this context. This research provides a comparative analysis of Naïve Bayes and SVM algorithms for sentiment classification related to unemployment in Indonesia during and after the pandemic.

The expected result of this analysis is to provide insights into the sentiments regarding unemployment in Indonesia and public sentiment trends. These insights can aid policymakers in addressing unemployment and improving welfare. Based on previous research, Naïve Bayes and SVM have shown varying levels of accuracy in different contexts. This study hypothesizes that one of these models will demonstrate superior performance in accurately classifying unemployment-related sentiment, thereby providing a clearer understanding of public opinion during and after the COVID-19 pandemic.

2. METHODS

This research begins with collecting tweet data from social media platform X (Twitter). Subsequently, the X (Twitter) data is divided into two sets: labeled data, which undergoes manual labeling, and unlabeled data. Next is the pre-processing stage, aimed at cleaning the data. Following this, the clean labeled data is tested and compared using the Naïve Bayes and Support Vector Machine to identify the algorithm with the highest accuracy. The next step involves labeling the unlabeled data with the model to classify the unlabeled data using the best algorithm, followed by creating visualizations based on the classification results. The proposed method to perform sentiment analysis on unemployment using X (Twitter) data is shown in Figure 1.

2.1. Data Collection

The tweet data is sourced from social media platform X (Twitter), collected through the Tweet-Harvest process using the keyword "pengangguran". Data collection through Tweet-Harvest can be configured by specifying the data collection time range, the keywords used, and the maximum amount of data.

Journal of Information Systems and Informatics
Vol. 6, No. 2, June 2024
p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

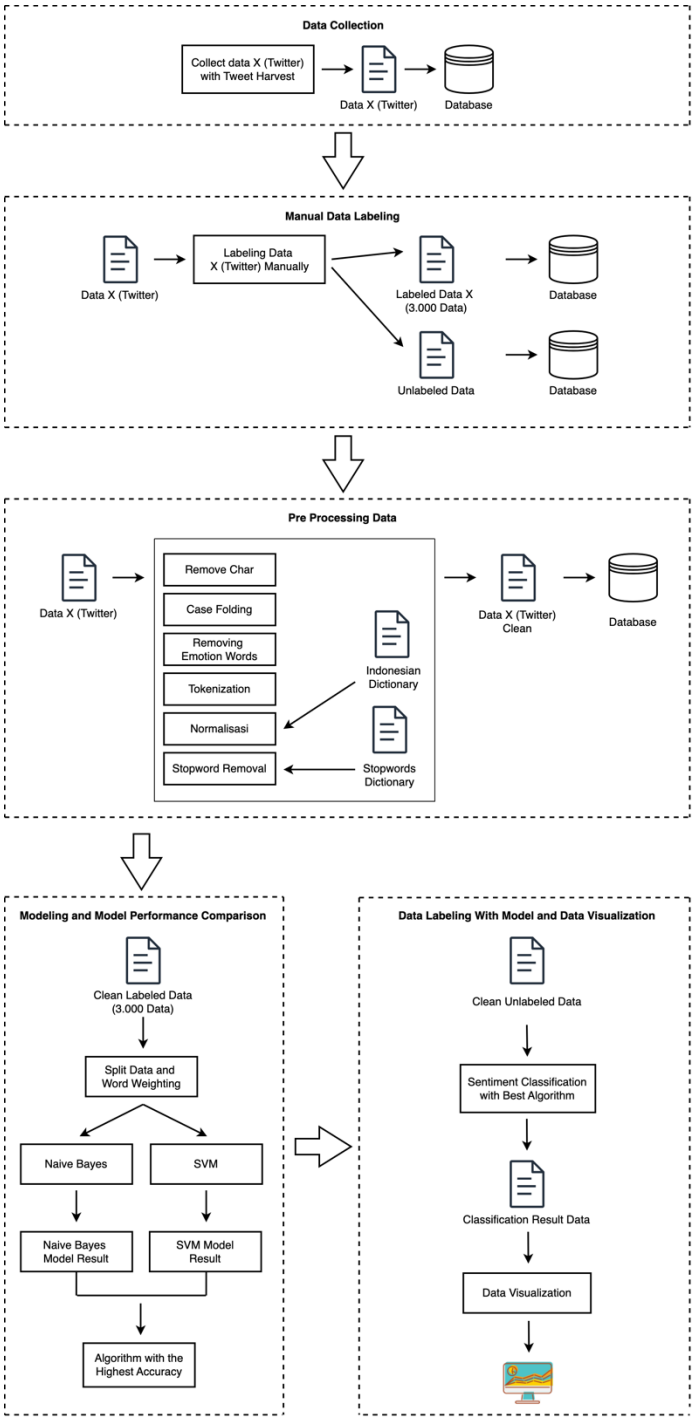

Figure 1. Research diagram

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

2.2. Manual Data Labeling

The label data consists of tweets that have been manually labeled and can be used to train models using Naïve Bayes and Support Vector Machine algorithms. In this study, three classes are used: "-1" indicating negative, "0" indicating neutral, and "1" indicating positive. The classes in the data are determined based on the context and words within the tweets. Negative tweets contain contexts that view unemployment negatively and refer to disappointment and dissatisfaction. They usually contain words such as "lazy," "bored," "burden." Positive tweets contain contexts that support unemployment. They typically contain words such as "spirit," "happy," "success." Meanwhile, in the neutral class, tweets do not express clear negative or positive feelings or views.

2.3. Pre-Processing Data

The preprocessing stage is the initial step in preparing and cleaning the data for the data modeling phase [10]. The following are the preprocessing steps in this study:

1) Remove character is a preprocessing step aimed at removing unnecessary characters such as hashtags, punctuation marks, emojis, URLs, and others.
2) Case folding is a conversion step where all characters/cases in document are converted to lowercase before modeling is performed [10].
3) Removing emotion words: In tweets, there are many variations of emotion words that are not important and are repeatedly written. These words need to be removed to help sentiment analysis models focus more on the main sentiment in the tweet.
4) Tokenization is a step involving the separation of text corpus into sentences that serve as the first-level tokens in the corpus.
5) Normalization is the process of transforming reviews or comments, including typo correction, abbreviation unification, and converting non standard language into standard language according to a normalization dictionary.
6) Stopword removal is a stage involving the selection of relevant words and the removal of words that do not have a significant impact on the document's content. For example, words like "*dan*", "*adalah*", "*di*," and "*yang*" [10].

2.4. Modeling

Before forming the model, word weighting using TF-IDF is performed. TF-IDF is a statistical approach that integrates both TF and IDF techniques to determine the relevance of a word within a document relative to a collection of documents

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

[11]. The more often the word appears, the higher its weight value. The following equation shows the TF-IDF formula.

$$TF - IDF\ (term, document) = TF(term, document) \times IDF(term) \quad (1)$$

Term Frequency (TF) quantifies the occurrence of a term within a document, while Inverse Document Frequency (IDF) assesses the significance of a term [11]. TF and IDF are then multiplied to obtain the TF-IDF value. After going through the word weighting stage, the model is formed using the Naïve Bayes and SVM algorithms. The Naïve Bayes method is a simple technique in probability classification where probabilities are calculated by summing the frequency and combination values of a dataset. The main advantage of using the Naïve Bayes classification method is that it requires a relatively small amount of training data in the text classification [12]. The Naïve Bayes model applied is of the Multinomial Naïve Bayes (MNB) category, commonly utilized in text analysis scenarios where data is depicted in word frequency vector format [13].

Support Vector Machine (SVM) is a supervised learning algorithm commonly applied in both classification and regression tasks. It functions by identifying a hyperplane that best separates two classes, optimizing the margin between them. The optimal hyperplane is situated equidistantly between the classes to maximize separation [11]. In SVM, a specified kernel function is employed to transform the original dimensions (lower dimensions) of the dataset into new dimensions (higher dimensions). To address the classification of non-linear data, adjustments need to be made by incorporating kernel functions. A kernel is a defining function for transforming the mapping of input data space into a new vector space with higher dimensions. This enables the separation of both classes using hyperplane lines in the new vector space [14]. In this study, the RBF kernel is used. The Radial Basis Function (RBF) kernel is often utilized because it generally delivers accurate results and is versatile enough to handle various types of data [11].

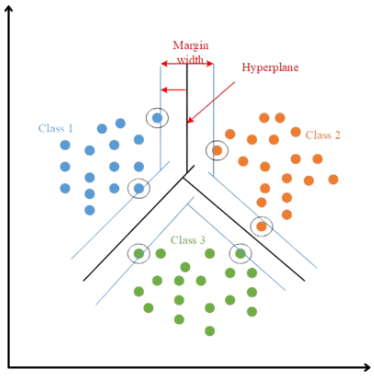

Figure 2. SVM Algorithm

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

Figure 2 illustrates the SVM algorithm. The blue, green, and orange points indicate three different classes. The margin represents the distance between the hyperplane and the nearest data points from each different class. Support vectors are the data points in each [15].

3. RESULTS AND DISCUSSION

3.1 Results

Tweets were collected from the website X (Twitter) with a time range from April 2020 to December 2023. The collected tweets will be stored in JSON format files and inserted into the database. The data obtained from the collecting process amounted to 578,854 data points. After being collected, 3000 data were manually labeled. Comprising 1000 data labeled -1, 1000 data labeled 0, and 1000 data labeled 1. The labeled data were then used to build the model. An example of tweet data and labeled data can be seen in Table 1.

Table 1. Tweet data and labeled data

No.	Tweet (Indonesia)	Tweet (English)	Class
1.	3000 pekerja coy di pecat. Klw numpuk 3 bulan di satu tempat pengangguran. Chaos !!!	3000 workers fired. If unemployed for 3 months in one place. Chaos !!!	-1
2.	Aneh bgt si para pengangguran ini	These unemployed people are really weird	-1
3.	pelajar,mahasiswa,pengangguran? mau penghasilan tambahan seperti itu? cepat hub saya sms : 083870782254 Pin : 2B613A4B https://t.co/OQeFVi32I6	students, college students, unemployed? Want extra income like that? Quickly contact me via SMS: 083870782254 Pin: 2B613A4B https://t.co/OQeFVi32I6	0
4.	@huftlepuff kalo pengangguran bisa sukses mah gamasalah	@huftlepuff if the unemployed can succeed, then there's no problem	1
5.	RUU Cipta Kerja merupakan terobosan utk mengurangi angka pengangguran. #DukungOmnibusLaw #OmnibusLawCiptaker https://t.co/O8xPV4Xhue	The Omnibus Law is a breakthrough to reduce unemployment rates. #SupportOmnibusLaw #OmnibusLawCiptaker https://t.co/O8xPV4Xhue	1

In this study, two algorithms were employed, namely Naïve Bayes and Support Vector Machine. There were four scenarios to test both algorithms, namely 90:10, 80:20, 70:30, and 60:40. The comparison results between the Naïve Bayes and Support Vector Machine algorithms is shown in Table 2.

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

Table 2. Comparison of classification models

Skenario	Algoritma	Accuracy	Class	Precision	Recall	F1-Score
90:10	Naïve Bayes	80.67%	-1	78.33%	**83.93%**	81.03%
			0	**92.59%**	76.53%	**83.80%**
			1	73.74%	81.11%	77.25%
	SVM	81.00%	-1	73.19%	**90.18%**	80.80%
			0	**98.65%**	74.49%	**84.88%**
			1	78.41%	76.67%	77.53%
80:20	Naïve Bayes	**81.50%**	-1	78.60%	**84.51%**	81.45%
			0	**90.70%**	78.39%	**84.10%**
			1	76.88%	81.38%	79.07%
	SVM	80.17%	-1	70.18%	**90.61%**	79.10%
			0	**96.64%**	72.36%	**82.76%**
			1	81.82%	76.60%	79.12%
70:30	Naïve Bayes	80.11%	-1	75.37%	**82.74%**	78.88%
			0	**86.96%**	77.67%	**82.05%**
			1	79.09%	79.93%	79.51%
	SVM	77.78%	-1	66.02%	**89.25%**	75.90%
			0	**94.02%**	71.20%	**81.03%**
			1	82.07%	72.54%	77.01%
60:40	Naïve Bayes	77.67%	-1	69.85%	**83.08%**	75.89%
			0	**86.34%**	76.33%	**81.03%**
			1	79.06%	73.59%	76.23%
	SVM	74.67%	-1	61.21%	**89.65%**	72.75%
			0	**93.44%**	68.84%	**79.28%**
			1	81.27%	65.64%	72.62%

This research emphasizes accuracy, which involves comparing the count of accurate predictions with the overall predictions made [16]. Accuracy provides an overview of the overall model performance. It measures how well the model can provide correct predictions in the corresponding categories [14]. The built model is then used to predict unlabeled data, as shown in Table 3.

Table 3. Predicted data

No	Tweet Clean (Indonesia)	Tweet Clean (English)	Class
1.	pusing pengangguran duluan	dizzy from being unemployed	-1
2.	pengangguran	unemployment	0
3.	pengangguran tidak enak lelah kerja lelah mencari kerja	unemployment is unpleasant, tired of working, tired of job hunting	-1
4.	terima kasih ayy resmi pengangguran	thank you, ayy, officially unemployed	1
5.	halah pengangguran	whatever, unemployment	-1

Journal of Information Systems and Informatics
Vol. 6, No. 2, June 2024
p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

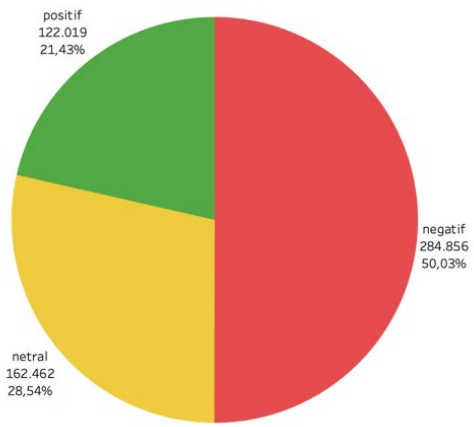

Figure 3. Sentiment percentage

Figure 3 shows the visualization of tweet data in percentage form. Based on the visualization, many Indonesians hold negative views on unemployment, totaling 50.03% or approximately 284,856 data. Furthermore, individuals who support unemployment amount to a total of 21.43% or 122,019 data. Meanwhile, 162,462 or 28.54% of Indonesians choose to remain neutral regarding the phenomenon of unemployment.

Figure 4. Sentiment trend

Figure 4 shows the visualization of the trend of public sentiment towards unemployment from April 2020 to December 2023. The visualization above shows that there are fluctuations in sentiment, which may be caused by various phenomena and factors, such as government policies, recession threats, as well as economic fluctuations.

3.2 Discussion

Based on Table 2, the Naïve Bayes algorithm achieved the highest accuracy score, which is 81.50%. This result was obtained from the 80:20 scenario, meaning 80% of the data for training and 20% for testing. The accuracy value of 81.50% indicates that the model can predict the data correctly for 81.50% of the total data used. Meanwhile, the remaining 18.50% of the data are predicted incorrectly by the model. Based on the conducted tests, this study employs the Naïve Bayes algorithm with the 80:20 scenario for the data labeling stage with the model. This is because the Naïve Bayes algorithm with the 80:20 scenario has the best accuracy value among the other scenarios and the Support Vector Machine algorithm. The accuracy value by the Naïve Bayes algorithm is higher than the SVM algorithm, which may be due to the data meeting the Bayes assumption, namely that the features are independent of each other. The existence of a characteristic within a particular category is independent of the presence of other characteristics [17]. SVM has limitations in handling data with similar features, which can impact accuracy values [18]. The recall value for the negative class tends to be higher, while the precision for the neutral class tends to be higher. However, there is no highest value for the positive class. This indicates that the model has poor performance in predicting positive data.

Based on Figure 4, October 2020 was the peak of negative sentiment, triggered by the enactment of the Omnibus Law policy, which led to protests and unrest. Demonstrations lasted for three days starting from October 6, 2020, in front of the West Java DPR building and Gedung Sate. This indicates significant public dissatisfaction and anxiety about potential job losses and economic instability caused by the new legislation. October 2022 was also the peak of negative sentiment, driven by the looming threat of a recession in 2023. The recession resulted in slowing trade, sluggish business activities, increased unemployment, rising in commodity prices, and decreasing purchasing power of the public. This reflects a widespread concern about job security and economic hardship, which can lead to psychological stress and anxiety among the unemployed and those fearing job loss. Meanwhile, April 2020 was the peak of positive sentiment. Despite being the early stages of the pandemic, the public held a positive view of unemployment as the government took steps to address it. One of which was President Joko Widodo's launch of social assistance for those impacted by the Coronavirus. This shows that government intervention can temporarily alleviate concerns and create a sense of hope and security among the unemployed. July

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

2023 was also the peak of positive sentiment. Unemployment rates in several regions decreased, with Jakarta's economy growing by 3.43% and unemployment decreasing by 13,000 people, a result of improved macroeconomic indicators. This illustrates that efforts to boost economic growth positively impact job opportunities and can improve the psychological well-being of the population by reducing financial stress and enhancing job security.

Figure 5. Word cloud for negative sentiment

Figure 5 shows the words which appear most frequently in negative sentiment. Words such as "*lelah*", "*sibuk*", and "*bosan*" reflect the perception that unemployment can lead to physical and mental fatigue, as well as feelings of boredom or despair due to unsuccessful job searching. The word "bodoh" reflects a negative view of individuals who are unemployed, with unemployment suggesting a lack of intelligence or ability. This may indicate the presence of negative stigma or stereotypes about unemployment in society. The words "miskin" and "beban" reflect the perception that unemployment can lead to financial difficulties and be a burden for affected individuals or families, highlighting the social and economic challenges posed by unemployment.

Figure 6. Word cloud for positive sentiment

Figure 6 shows the words which appear most frequently in positive sentiment. The appearance of "*enak*", "*kaya*", "sukses", "*bahagia*" indicates that society sees unemployment as something enjoyable and relaxing, and even though

unemployed, one can still achieve wealth, success, and happiness. The words "selamat" and "semangat" indicate that with enthusiasm, effort, and support, one can overcome the challenges of unemployment. This can encourage discussions and actions that support solutions to reduce unemployment rates and increase job opportunities.

4. CONCLUSION

This study compares the Naïve Bayes and Support Vector Machine (SVM) algorithms with TF-IDF weighting for sentiment classification. The data used in the study were collected from social media platform X (Twitter) using Tweet-Harvest with the keyword "pengangguran". The total number of data obtained from April 2020 to December 2023 was 576,764 tweets. The highest accuracy rate was achieved with the Naïve Bayes algorithm at 81.50% with an 80:20 scenario (80% training data and 20% testing data). The sentiment analysis results obtained indicate that 50.03% of the data were classified as negative sentiment, 28.54% as neutral sentiment, and 21.43% as positive sentiment. Public sentiment trends are influenced by various phenomena and factors, such as government policies, recession threats, as well as economic fluctuations. Peaks in negative sentiment correlated with specific events such as the Omnibus Law enactment and recession threats, while positive sentiment peaks followed government support measures and economic recovery. These findings highlight the need for policymakers to consider public sentiment in addressing unemployment and improving welfare. This research contributes by offering comprehensive insights into public opinion on unemployment during and post-COVID-19, and by demonstrating the comparative effectiveness of Naïve Bayes and SVM for sentiment analysis.

REFERENCES

[1] S. Panneer *et al.*, "Health, Economic and Social Development Challenges of the COVID-19 Pandemic: Strategies for Multiple and Interconnected Issues," *Healthcare (Switzerland)*, vol. 10, no. 5, May 2022, doi: 10.3390/healthcare10050770.

[2] N. Aeni, "Pandemi COVID-19: Dampak Kesehatan, Ekonomi, dan Sosial COVID-19 Pandemic: The Health, Economic, and Social Effects," vol. 17, no. Juni, pp. 17–34, 2021, doi: https://doi.org/10.33658/jl.v17i1.249.

[3] F. Rizal and H. Mukaromah, "Kebijakan Pemerintah Indonesia Dalam Mengatasi Masalah Pengangguran Akibat Pandemi Covid 19," 2021.

[4] N. Rihhadatul'aisyi, S. Muthmainnah, T. W. Putri, H. P. Zahra, and F. T. Febrian, "Efek Twitter di Masa Pandemi COVID-19 pada Sikap dan Perilaku," *Jurnal Ilmu Komunikasi*, vol. 19, no. 2, p. 205, Oct. 2021, doi: 10.31315/jik.v19i2.4178.

[5] R. Satria, I. M. A. D. Suarjaya, and I. P. A. E. Pratama, "Sentiment Analisi Antusias Masyarakat Terhadap Sampah Plastik Dengan Menggunakan Metode Support Vector Machine (SVM)," *JITTER- Jurnal Ilmiah Teknologi dan Komputer*, vol. 3, no. 1, 2022.

[6] E. S. Romaito, M. K. Anam, R. Rahmaddeni, and A. N. Ulfah, "Perbandingan Algoritma Svm Dan Nbc Dalam Analisa Sentimen Pilkada Pada Twitter," *CSRID (Computer Science Research and Its Development Journal)*, vol. 13, no. 3, p. 169, Nov. 2021, doi: 10.22303/csrid.13.3.2021.169-179.

[7] K. M. A. Candrayani, I. M. A. D. Suarjaya, and A. A. K. A. C. Wiranatha, "Analisis Sentimen Pembelajaran Daring Era Pandemi COVID-19 Menggunakan Naive Bayes Dan SVM," *TEMATIK*, vol. 10, no. 1, pp. 47–53, Jun. 2023, doi: 10.38204/tematik.v10i1.1274.

[8] M. Rizal Ramli and H. Sulastri, "Sentiment Analysis Of Student Opinion Related To Online Learning Using Naïve Bayes Classifier Algorithm And SVM With Adaboost On Twitter Social Media," *Jurnal Informatika dan Teknologi Informasi*, vol. 20, no. 2, pp. 187–201, 2023, doi: 10.31515/telematika.v20i2.8827.

[9] M. Meenakshi and J. Kaur, "Analysis of Employment-Related Sentiment in Social Media in India," *Mathematical Statistician and Engineering Applications*, vol. 71, no. 4, pp. 10277–10288, 2022.

[10] B. W. Kurniadi, N. H. Matondang, and D. S. Prasvita, "Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Penggunaan Aplikasi Jobstreet Implementation of Naive Bayes Algorithm For Sentiment Analysis Of The Use Of Jobstreet Application," *Agustus*, vol. 21, no. 3, pp. 523–533, 2022, doi: https://doi.org/10.33633/tc.v21i3.6361.

[11] P. H. Prastyo, I. Ardiyanto, and R. Hidayat, "Indonesian Sentiment Analysis: An Experimental Study of Four Kernel Functions on SVM Algorithm with TF-IDF," in *2020 International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy, ICDABI 2020*, Institute of Electrical and Electronics Engineers Inc., Oct. 2020. doi: 10.1109/ICDABI51230.2020.9325685.

[12] P. S. M. Suryani, L. Linawati, and K. O. Saputra, "Penggunaan Metode Naïve Bayes Classifier pada Analisis Sentimen Facebook Berbahasa Indonesia," *Majalah Ilmiah Teknologi Elektro*, vol. 18, no. 1, p. 145, May 2019, doi: 10.24843/mite.2019.v18i01.p22.

[13] S. Khomsah, "Naive Bayes Classifier Optimization on Sentiment Analysis of Hotel Reviews," *Jurnal Penelitian Pos dan Informatika*, vol. 10, no. 2, pp. 157–168, Dec. 2020, doi: 10.17933/jppi.2020.100206.

[14] M. A. S. addam and E. D. Kurniawan, "Analisis Sentimen Fenomena PHK Massal Menggunakan Naive Bayes dan Support Vector Machine," *Jurnal Informatika: Jurnal pengembangan IT (JPIT)*, vol. 8, no. 3, pp. 226–233, 2023, doi: https://doi.org/10.30591/jpit.v8i3.4884.

[15] E. Cengil and A. Çınar, "The effect of deep feature concatenation in the classification problem: An approach on COVID-19 disease detection," *Int J Imaging Syst Technol*, vol. 32, no. 1, pp. 26–40, Sep. 2021, doi: 10.1002/ima.22659.

[16] S. N. J. Fitriyyah, N. Safriadi, and E. E. Pratama, "Analisis Sentimen Calon Presiden Indonesia 2019 dari Media Sosial Twitter Menggunakan Metode Naive Bayes," *JEPIN (Jurnal Edukasi dan Penelitian Informatika)* , vol. 5, no. 3, pp. 279–285, 2019, doi: http://dx.doi.org/10.26418/jp.v5i3.34368.

[17] T. A. Azzahra *et al.*, "Perbandingan Efektivitas Naïve Bayes dan SVM dalam Menganalisis Sentimen Kebencanaan di Youtube," *Jurnal Media Informatika Budidarma*, vol. 8, no. 1, pp. 312–322, 2024, doi: 10.30865/mib.v8i1.7186.

[18] S. I. Nurhafida and F. Sembiring, "Analisis Sentimen Aplikasi Novel Online Di Google Play Store Menggunakan Algoritma Support Vector Machine (SVM)," *Jurnal Sains Komputer & Informatika (J-SAKTI)*, vol. 6, no. 1, pp. 317–327, 2022.