Mining Public Sentiment and Trends in Social Media Discussions on Indonesian Presidential Candidates Using Support Vector Machines

Satrya Fajri Pratama[1], Arif Mu'amar Wahid[2,*]

[1]Department of Computer Science, School of Physics, Engineering and Computer Science, University of Hertfordshire, United Kingdom

[2]Master of Computer Science, Computer Science Faculty, Universitas Amikom Purwokerto, Indonesia

ABSTRACT

This study investigates public sentiment surrounding the 2024 Indonesian presidential candidates by analyzing Twitter data collected between October 2022 and April 2023. Using Support Vector Machines (SVM) for sentiment classification, the research aims to uncover patterns in early public opinion during the pre-election period. A dataset of 28,782 tweets mentioning three main candidates—Anies Baswedan, Ganjar Pranowo, and Prabowo Subianto—was collected, cleaned, and processed. Exploratory Data Analysis (EDA) revealed that positive sentiment dominated discussions for all candidates, with Ganjar Pranowo receiving the highest proportion of positive tweets. Geographic and user engagement metrics further contextualized the digital discourse, highlighting differences in audience demographics and regional engagement. The text data was transformed using Term Frequency-Inverse Document Frequency (TF-IDF) vectorization, capturing both unigrams and bigrams, and limited to 5,000 features for computational efficiency. The dataset was split into training and testing sets with a 75/25 ratio. A linear kernel SVM was trained with class weighting to address class imbalance, achieving an accuracy of 87.9% on the test set. Precision, recall, and F1-score metrics indicated strong model performance, particularly for the positive sentiment class. Feature importance analysis identified key terms influencing classification, aligning intuitively with sentiment polarity. These findings demonstrate the viability of using SVM and social media data to analyze early political sentiment in emerging democratic contexts. The study contributes to the growing field of digital political analytics by providing insights into voter sentiment dynamics prior to official candidacy announcements. While limited to Twitter and a specific timeframe, the research lays groundwork for multi-platform, longitudinal studies. Future research could explore sentiment evolution throughout the election cycle and deeper analysis of sentiment drivers by region and policy.

Keywords Sentiment Analysis, Support Vector Machines, Indonesian Presidential Election, Social Media Mining, Political Discourse

Submitted 26 May 2025
Accepted 1 June 2025
Published 2 June 2025

*Corresponding author
Arif Mu'amar Wahid,
arif@amikompurwokerto.ac.id

Additional Information and Declarations can be found on page 149

Introduction

The role of social media in political campaigns has become increasingly significant, reshaping public opinion and influencing election outcomes globally. This transformation is primarily driven by the ability of social media platforms to facilitate direct communication between political actors and the electorate, allowing for a more personalized and interactive form of political engagement. Research indicates that political communication in the age of social media has evolved into a critical component of election strategies, as politicians leverage

How to cite this article: S.F.Pratama and A.M.Wahid, "Mining Public Sentiment and Trends in Social Media Discussions on Indonesian Presidential Candidates Using Support Vector Machines," *J. Digit. Soc.*, vol. 1, no. 2, pp. 138-151, 2025.

these platforms to disseminate their messages and engage with voters [1], [2], [3].

Social media provides a unique space for political actors to craft and share their narratives, often tailored to resonate with specific demographics. For instance, the integration of social media into political campaigns has been shown to enhance the effectiveness of communication strategies, as evidenced by the synchronized activities of politicians across various platforms [4], [5]. This hybrid media system allows for a more nuanced approach to political messaging, where traditional media outlets are complemented by the immediacy and accessibility of social media. The ability to interact with constituents in real-time fosters a sense of community and engagement, which is crucial for mobilizing support during elections [6].

Moreover, the impact of social media on political participation cannot be overstated. Studies have demonstrated that exposure to political content on social media platforms significantly increases civic engagement among users, encouraging them to participate in political discussions and activities [7], [8]. This trend is particularly pronounced among younger demographics, who often rely on social media as their primary source of political information. As such, political campaigns that effectively utilize social media can enhance voter turnout and engagement, as seen in various electoral contexts worldwide [9], [10].

The examination of Twitter discussions surrounding the 2024 Indonesian Presidential candidates is essential for understanding the evolution of public sentiment over time. Social media platforms, particularly Twitter, have become pivotal in shaping political discourse and public opinion, providing a real-time reflection of voter sentiments and preferences. This study aims to analyze the sentiments expressed in tweets related to the presidential candidates, utilizing sentiment analysis techniques to quantify and interpret public opinion.

Research indicates that Twitter serves as a significant platform for political discussions in Indonesia, where users engage in more deliberate and civil discourse compared to other platforms [11]. The ability to analyze sentiments on Twitter allows researchers to track changes in public opinion as the election approaches, providing insights into how candidates are perceived by the electorate. For instance, sentiment analysis has been successfully applied in previous studies to gauge public reactions to political events and candidates, demonstrating its effectiveness in predicting electoral outcomes [12], [13].

The methodology for sentiment analysis in this context involves the use of natural language processing (NLP) techniques to classify the sentiments expressed in tweets. Tools such as the VADER (Valence Aware Dictionary and sEntiment Reasoner) algorithm have been employed to assess the polarity of tweets, categorizing them into positive, negative, or neutral sentiments [14]. This approach allows for a nuanced understanding of public sentiment, revealing not only the overall favorability of candidates but also the intensity of public emotions towards them [15].

The primary objective of this study is to analyze public sentiment related to the three main presidential candidates for the 2024 Indonesian election by mining Twitter data collected between October 2022 and April 2023. Utilizing Support Vector Machines (SVM), a powerful machine learning algorithm for text classification, the research aims to accurately classify the sentiment of tweets as positive, negative, or neutral, thereby revealing patterns and trends in public opinion during this pre-candidacy period.

This study holds considerable significance as it sheds light on early-stage voter sentiment in a developing democratic environment, offering insights that can assist political stakeholders and analysts. Moreover, it demonstrates the effectiveness of SVM in sentiment analysis of social media data, providing a methodological contribution that can be applied in future political data mining research within Indonesia and similar contexts.

Literature Review

Sentiment Analysis in Social Media

Sentiment analysis on social media platforms, particularly Twitter, has emerged as a vital tool for gauging public opinion during elections. Numerous studies have highlighted the effectiveness of sentiment analysis in understanding voter sentiments and preferences, especially in the context of short, informal tweets. The challenge of sentiment classification is well-documented, as the unique characteristics of tweets—such as brevity, informal language, and the use of slang—can complicate the sentiment analysis process [16].

One of the primary challenges in sentiment classification is the informal nature of tweets, which often includes misspellings, abbreviations, and colloquial expressions. These linguistic features can hinder the accuracy of automated sentiment analysis tools, necessitating advanced preprocessing techniques to clean and normalize the data [17]. For instance, a study by Alvi et al. proposed a recursive text preprocessing approach that enhances the quality of the dataset by removing noisy data and irrelevant tweets, thereby improving sentiment classification outcomes [18]. This highlights the importance of developing robust methodologies that can effectively handle the informal structure of tweets.

Furthermore, the integration of aspect-based sentiment analysis allows for a more nuanced understanding of public opinion by focusing on specific topics or attributes related to candidates or policies. This approach can reveal not only the overall sentiment towards a candidate but also the sentiments associated with particular issues, such as economic policies or social justice [19]. Such detailed insights can significantly inform political campaigns and strategies, enabling candidates to address specific voter concerns more effectively.

Support Vector Machines for Text Classification

Support Vector Machines (SVM) have gained prominence in text classification tasks due to their ability to handle high-dimensional data and effectively separate classes. The foundational work by Cortes and Vapnik in 1995 established SVM as a robust supervised learning method that excels in various applications, including sentiment analysis [20]. SVM's effectiveness in sentiment analysis is attributed to its capacity to create hyperplanes that optimally differentiate between classes, making it particularly suitable for the nuanced nature of textual data.

Recent studies have demonstrated the successful application of SVM in sentiment analysis across different contexts. For instance, a study by Sester et al. highlighted that SVM-based models achieved competitive performance in classifying sentiments, particularly when dealing with complex and non-linear relationships between textual features and sentiment labels [21]. This adaptability is crucial in sentiment analysis, where the relationships between words and sentiments can be intricate and context-dependent.

Moreover, Aditiya et al. utilized SVM for sentiment analysis of user reviews on the MyIM3 application, employing a confusion matrix to evaluate the model's performance. Their findings underscored the effectiveness of SVM in accurately classifying sentiments, reinforcing its utility in practical applications [22]. Similarly, Imanuddin reported high accuracy rates (80-90%) for SVM in sentiment analysis tasks, emphasizing its reliability and ease of implementation [23].

Gaps in Existing Literature

The existing literature on sentiment analysis in the context of Indonesian politics reveals a significant gap regarding the analysis of early-stage, pre-candidacy sentiment on social media. While there have been studies focusing on social media sentiment during election periods, few have specifically targeted the pre-candidacy phase, which is crucial for understanding the evolving public perception of potential candidates. This study aims to address this gap by applying Support Vector Machines (SVM) to sentiment analysis over time for the 2024 presidential election in Indonesia.

Research by Fitriani highlights the importance of analyzing social media sentiments within the Indonesian context, particularly how these sentiments relate to macroeconomic indicators [24]. Although this study does not focus on political candidates, it underscores the relevance of sentiment analysis in understanding public opinion in Indonesia, suggesting that similar methodologies could be applied to political contexts, especially during the pre-candidacy phase.

Conrad et al. discuss the potential of social media as an alternative to traditional surveys for gauging public opinion about economic conditions [25]. This perspective can be extended to political sentiment, indicating that social media platforms like Twitter can provide valuable insights into public sentiment regarding potential presidential candidates before they officially declare their candidacy. The authors emphasize the differences in data generation between surveys and social media, which could be critical in analyzing early-stage sentiments.

While some studies have employed sentiment analysis in Indonesian politics, such as the work by Utomo et al. on clustering data based on social media [26], they often focus on specific events or candidates rather than the broader pre-candidacy sentiment landscape. This highlights the need for a comprehensive analysis that captures the nuances of public sentiment as potential candidates emerge.

Method

This study follows a systematic pipeline to analyze Twitter data on the 2024 Indonesian presidential candidates using Support Vector Machines (SVM) for sentiment classification. The methodology includes data integration, cleaning, exploratory data analysis (EDA), text preprocessing, feature extraction, model training with parameter tuning, evaluation, and interpretation.

Data Collection and Integration

The dataset comprises three CSV files containing tweets related to each candidate: Anies Baswedan, Ganjar Pranowo, and Prabowo Subianto. These files were loaded and concatenated into a single dataframe, with an added column identifying the candidate for each tweet to enable comparative analysis. During integration, extraneous index columns such as 'Unnamed: 0' or 'Unnamed: 0.1', commonly generated by CSV exports, were removed to avoid redundant data and potential misalignment. This ensures a clean unified dataset spanning the period from October 2022 to April 2023.

Data Cleaning

Initial cleaning involved standardizing column names and converting date-related columns ('Date' and 'Created') into datetime objects for accurate temporal analysis. Rows lacking essential information such as tweet text, sentiment labels, or date were removed to maintain dataset quality. Tweets were filtered to retain only those with sentiment labels ‘Positive’ or ‘Negative’ to establish a binary classification framework. Sentiment labels were then encoded numerically: ‘Positive’ as 1 and ‘Negative’ as 0. Location data underwent normalization by converting text to lowercase and trimming spaces; missing or unspecified locations were replaced with the placeholder 'not specified'. This preprocessing ensures consistent, high-quality input for analysis and modeling.

Exploratory Data Analysis (EDA)

EDA was conducted to understand the underlying distribution and trends within the dataset. Sentiment distribution was examined by aggregating counts and percentages per candidate, visualized through bar charts to compare the volume of positive and negative tweets. The daily tweet volume per candidate was plotted as a time series, resampling tweets by day, revealing fluctuations and potential correlations with political events. User-related metrics such as average number of followers, following, and total tweets per candidate’s tweet authors were calculated to assess user influence and engagement. These metrics were visualized with grouped bar charts for comparison. Additionally, the geographic distribution of tweets was analyzed for locations explicitly specified by users, with the top 10 locations per candidate highlighted. Word clouds were generated after applying custom stopwords—combining standard English and Indonesian stopwords with domain-specific terms like candidate names and common Twitter slang—to visually represent the most frequent words in positive, negative, and all tweets.

Text Preprocessing for Modeling

To prepare textual data for machine learning, tweets were preprocessed by converting to lowercase and removing noise such as URLs, mentions (user handles), hashtags (while preserving the word itself), punctuation, and numerical characters. Stopwords from both English and Indonesian languages were removed using NLTK’s stopword corpus supplemented by custom stopwords to exclude candidate names and frequent Twitter shorthand (e.g., ‘rt’, ‘yg’). The preprocessing function also discarded words shorter than two characters to eliminate tokens unlikely to contribute meaningful information. This step is crucial to reduce dimensionality and noise, improving model performance.

Feature Extraction with TF-IDF

Text data was transformed into numerical feature vectors using Term

Frequency-Inverse Document Frequency (TF-IDF) vectorization, a method that weighs terms by their frequency in individual documents relative to their frequency across the entire corpus. This highlights words that are distinctive for a tweet while down-weighting common terms. The vectorizer was configured to extract both unigrams and bigrams (sequences of one or two words), capturing contextual information beyond single words. A maximum feature limit of 5,000 was imposed to balance computational efficiency and feature richness. This parameter helps prevent overfitting and reduces memory usage during model training.

Model Training with Support Vector Machines

The dataset was split into training (75%) and testing (25%) subsets using stratified sampling to maintain the class distribution in both sets, ensuring that both positive and negative sentiment tweets were proportionally represented. The Support Vector Machine model was trained using a linear kernel, selected for its efficiency and interpretability in high-dimensional text classification. The model was implemented with the LinearSVC class from scikit-learn, configured with a regularization parameter C=1.0 which controls the trade-off between maximizing the margin and minimizing classification error. The class_weight='balanced' option was used to automatically adjust weights inversely proportional to class frequencies, addressing any imbalance between positive and negative samples. A maximum iteration limit of 5000 was set to ensure convergence, while the random_state parameter fixed to 42 guarantees reproducibility of results.

Model Evaluation and Interpretation

After training, the model was evaluated on the test set using metrics such as accuracy (overall correctness), precision (correct positive predictions out of total positive predictions), recall (correct positive predictions out of total actual positives), and F1-score (harmonic mean of precision and recall). A confusion matrix was plotted to visually summarize true positives, true negatives, false positives, and false negatives, providing insights into classification errors. For interpretability, feature importance was extracted from the SVM's coefficient vector, which reflects the influence of each TF-IDF feature (word or n-gram) on classification decisions. The top 20 features with the highest positive coefficients were identified as indicators of positive sentiment, while the 20 lowest (most negative) coefficients corresponded to features signaling negative sentiment. This analysis helps understand which words most strongly influence the model's sentiment predictions.

Result and Discussion

Data Loading and Cleaning

The initial dataset consisted of 30,000 tweets collected across three candidates: Anies Baswedan, Ganjar Pranowo, and Prabowo Subianto. After cleaning steps such as removing rows with missing essential data and filtering for relevant sentiment labels, the dataset was reduced to 28,782 records. This ensured high data quality by excluding incomplete entries and non-binary sentiment labels, providing a solid foundation for analysis.

Exploratory Data Analysis (EDA)

The sentiment distribution in figure 1 revealed that all three candidates attracted more positive than negative tweets. Ganjar Pranowo had the highest proportion of positive sentiment (79.15%), followed by Prabowo Subianto (74.46%), and Anies Baswedan (65.08%). In absolute terms, Ganjar also led in positive tweet counts, with 7,559 positive tweets out of 9,550 total tweets. These results suggest relatively strong positive public engagement for Ganjar and Prabowo during the pre-election period.

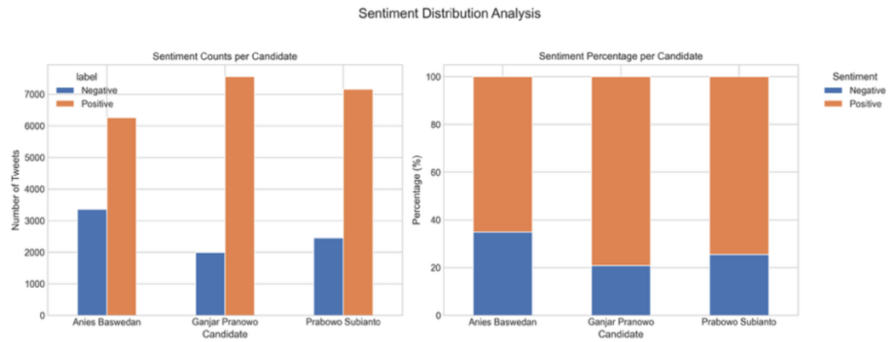

Figure 1 **Sentiment Distribution**

Analysis of tweet volume over time (figure 2) showed similar average daily tweet counts across candidates (around 55 tweets/day), but with significant variability for Anies Baswedan and Prabowo Subianto (standard deviations above 55), indicating fluctuating public attention. Ganjar's tweet volume was more consistent, with a smaller standard deviation (~23), and fewer minimum daily tweets (14) compared to others.

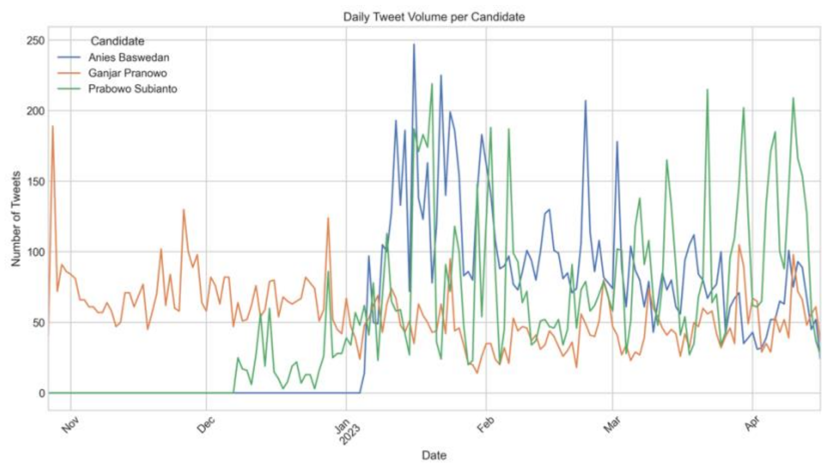

Figure 2 **Daily Tweet Volume**

User metrics (figure 3) highlighted distinct differences in the typical Twitter user engaging with each candidate. Anies Baswedan's followers averaged around 97,000 per user, while Ganjar's averaged roughly 79,700 and Prabowo's 36,500. Ganjar's followers also tended to follow more accounts on average (about 1,083), suggesting different social network behaviors. Prabowo's users had notably fewer followers and tweet counts, potentially indicating different audience segments or engagement patterns.

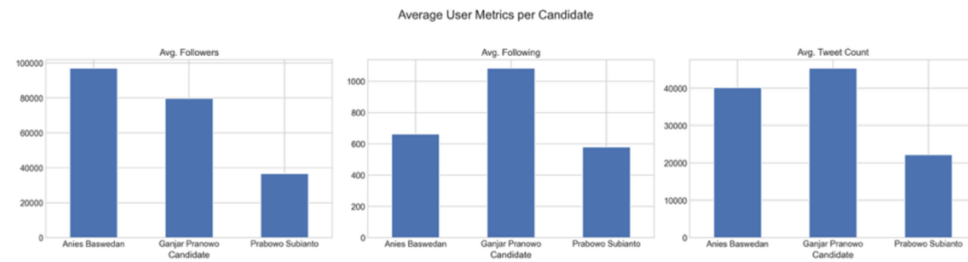

Figure 3 **Average User Metrics**

Geographic analysis of tweet locations found that most tweets originated from major urban areas such as Jakarta (across various administrative districts), Indonesia as a whole, and regions like Banten, East Nusa Tenggara, and Central Java. Notably, Anies Baswedan's tweets were heavily concentrated in Jakarta and surrounding areas, while Ganjar's and Prabowo's audiences appeared more geographically dispersed. These spatial insights may reflect candidate-specific voter bases or regional campaign strengths.

Word clouds created for all tweets (figure 4), as well as separately for positive and negative sentiments, highlighted key themes and vocabulary. Common positive words included "support," "reward," "best," "hope," and "win," while negative sentiment was associated with terms such as "punishment," "death," "stupid," "violence," and "fail." These thematic differences align with expected sentiment polarities and provide qualitative context to the quantitative findings.

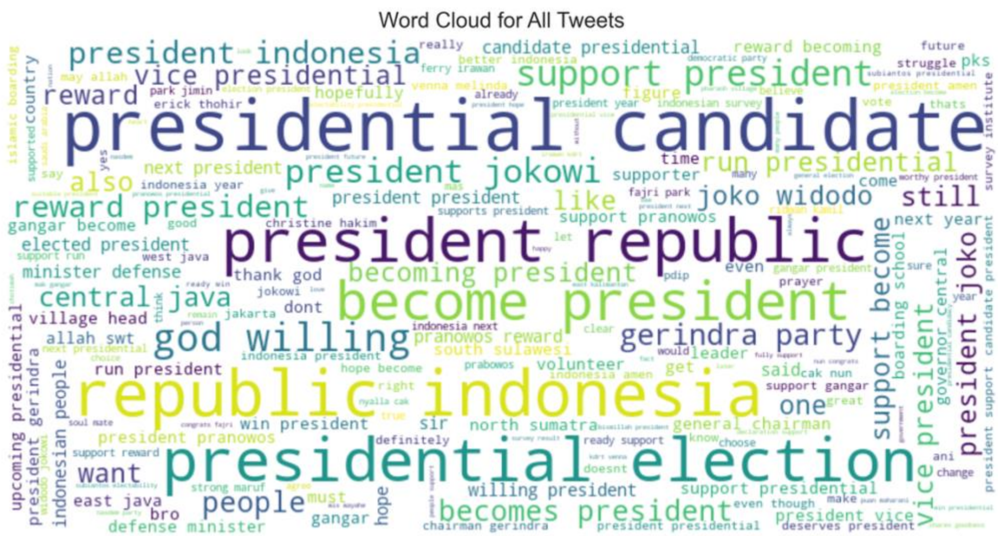

Figure 4 **Word Cloud**

Model Training and Evaluation

The dataset was split into training (21,586 tweets) and testing (7,196 tweets) subsets while maintaining sentiment class proportions. TF-IDF vectorization was applied with a limit of 5,000 features and inclusion of both unigrams and bigrams, capturing important lexical and contextual information from tweets. The vectorizer was saved for reproducibility. A linear kernel Support Vector Machine (LinearSVC) model was trained with class-weight balancing and a regularization parameter C=1.0. Training was efficient, completing in just 0.34 seconds. The trained model was saved for future use. On the test set, the SVM model achieved an accuracy of approximately 87.94%, demonstrating robust performance in distinguishing positive and negative sentiments. The classification report showed strong recall and precision for the positive class (recall 88%, precision 95%) and somewhat lower but still good scores for the negative class (recall 87%, precision 73%), indicating the model was particularly effective at identifying positive tweets. The confusion matrix (figure 5) further illustrated model performance, with 1,702 true negatives and 4,626 true positives, alongside 250 false positives and 618 false negatives. This suggests that while the model is slightly more prone to misclassify some negative tweets as positive, overall classification quality is high.

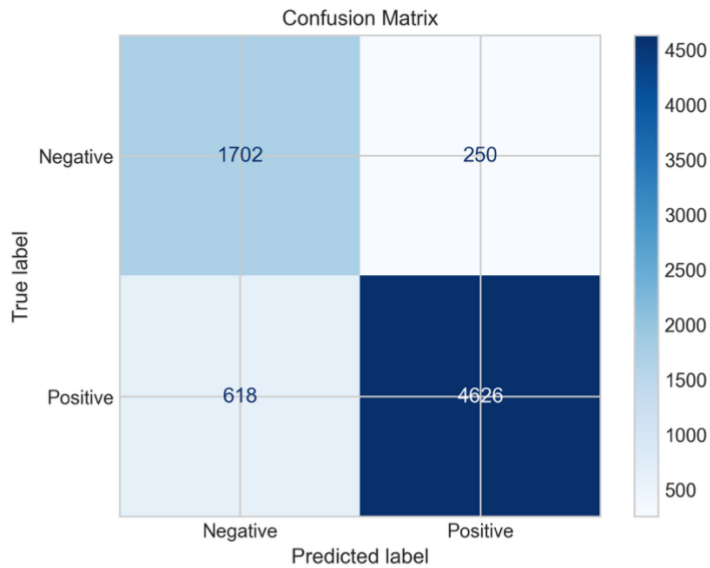

Figure 5 Confusion Matrix

Feature Importance Analysis

The SVM's linear coefficients were examined to identify the top words influencing sentiment classification (figure 6). Among the top 20 features contributing positively were words like “support,” “reward,” “best,” “hope,” and “win,” which align intuitively with positive public opinion. Conversely, negatively weighted features included “punishment,” “death,” “stupid,” “violence,” and “fail,” reflecting terms commonly associated with negative sentiment. These findings provide interpretability to the model, showing that it captures meaningful sentiment indicators relevant to the political discourse on Twitter. The feature importance visualization clearly delineates words that drive sentiment classification, supporting confidence in the model's decision-making process.

The results indicate a generally positive tone in the Twitter discourse around all three presidential candidates during the pre-election period, with Ganjar Pranowo receiving the highest positive sentiment proportion. The SVM model demonstrated strong classification performance, confirming its suitability for sentiment mining in this political context. EDA insights on user metrics and geographic distribution add depth to the understanding of the electorate's digital behavior, while feature importance sheds light on the linguistic markers shaping public sentiment. These outcomes lay a strong foundation for further research on social media influence in Indonesian politics.

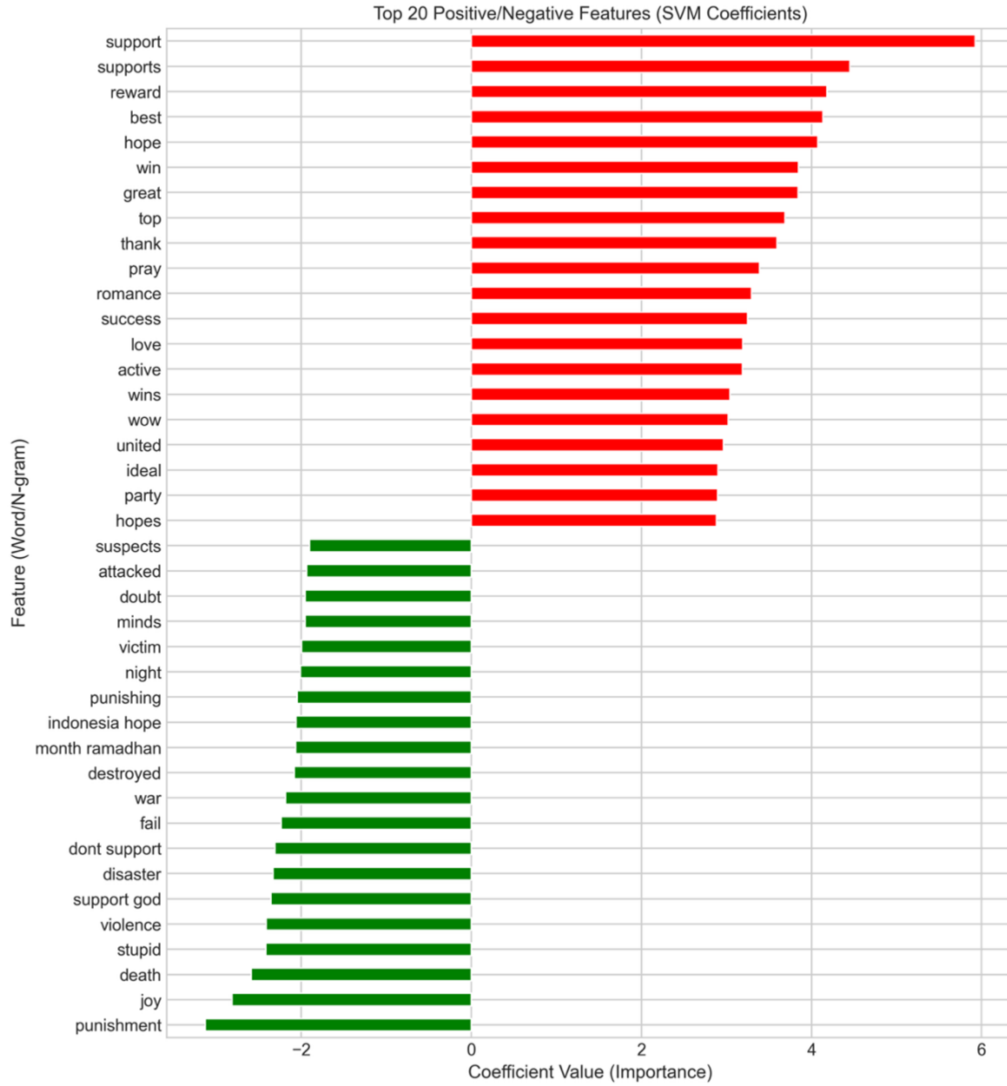

Figure 6 Top 20 Feature

Discussion

The Support Vector Machine (SVM) model applied in this study proved to be effective in classifying public sentiment from Twitter data, despite the inherent challenges posed by informal language, slang, abbreviations, and noise commonly found in social media text. The high accuracy and strong precision and recall scores demonstrate that SVM, combined with careful preprocessing and TF-IDF vectorization, can robustly handle the complexity of short, noisy tweets. This validates the suitability of machine learning techniques for analyzing public opinion in the dynamic and informal context of social media discussions.

The sentiment analysis results reveal significant early public engagement and interest in the 2024 Indonesian presidential election, with a generally positive tone prevailing across the three main candidates. The differing levels of positive and negative sentiment among candidates suggest variations in their public appeal and highlight key themes and topics that resonated with voters during the pre-campaign period. The geographic and user metric analyses further contextualize these sentiments, indicating how voter demographics and regional factors may shape digital political discourse.

These findings have practical implications for political campaign teams and strategists. Understanding the sentiment landscape and the topics that generate strong positive or negative reactions can help tailor messaging, policy emphasis, and outreach efforts to better connect with the electorate. Early identification of public concerns or enthusiasm through social media sentiment mining allows campaigns to be more agile and responsive, potentially influencing voter perceptions before formal campaign activities intensify.

Beyond immediate campaign strategy, this research contributes to the broader understanding of how digital sentiment analysis can inform electoral studies. It highlights the potential for social media data and machine learning models like SVM to track evolving public opinion in real time, offering valuable insights to political scientists, sociologists, and data analysts. Future studies can build on this work by integrating multi-platform data, exploring causal links between sentiment and election outcomes, and refining sentiment models to capture nuanced emotions and contextual subtleties in political communication.

Conclusion

This study successfully applied Support Vector Machines (SVM) to analyze public sentiment toward the 2024 Indonesian presidential candidates using Twitter data. The results uncovered significant sentiment patterns, highlighting how positive and negative opinions varied among candidates during the pre-election period. This work demonstrated that machine learning techniques like SVM are effective tools for mining and interpreting political sentiment from large-scale social media datasets, providing valuable insights into early voter behavior and discourse.

The research contributes meaningfully to the expanding field of digital sentiment analysis in political science, particularly within emerging democracies where social media increasingly shapes public opinion. By showcasing the application of SVM to Indonesian political Twitter data, this study offers a methodological framework for analyzing informal and noisy social media text in politically dynamic contexts. It also underscores the potential of computational approaches to complement traditional survey methods in understanding electoral sentiment and public engagement.

However, the study has limitations. The analysis was restricted to Twitter as the sole data source, which may not capture the full spectrum of political discussions occurring on other popular platforms such as Facebook or Instagram. Additionally, the dataset covers a limited time frame before the official candidacy declarations, providing only a snapshot of evolving sentiment. These constraints suggest the findings may not fully represent the broader public opinion or sentiment dynamics over the entire election cycle.

Future research could expand by incorporating data from multiple social media platforms to gain a more holistic view of political sentiment. Longitudinal studies tracking sentiment changes throughout the entire election process, including post-election periods, would enrich understanding of voter behavior and campaign impact. Moreover, deeper sentiment analysis exploring regional variations or linking sentiment to specific candidate policies could uncover the underlying drivers of public opinion shifts, enhancing the practical value of sentiment mining in political strategy and social research.

Declarations

Author Contributions

Conceptualization: S.F.P.; Methodology: A.M.W.; Software: S.F.P.; Validation: A.M.W.; Formal Analysis: A.M.W.; Investigation: S.F.P.; Resources: A.M.W.; Data Curation: S.F.P.; Writing Original Draft Preparation: A.M.W.; Writing Review and Editing: A.M.W.; Visualization: S.F.P.; All authors have read and agreed to the published version of the manuscript.

Data Availability Statement

The data presented in this study are available on request from the corresponding author.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] A. M. Hadma and J. D. Anggoro, “Political Communication in the Age of Social Media,” *Commicast*, vol. 3, no. 1, pp. 1–7, 2021, doi: 10.12928/commicast.v3i1.5114.

[2] D. Subekti, D. Mutiarin, and A. Nurmandi, “Political Communication in Social Media: A Bibliometrics Analysis,” *Stud. Media Commun.*, vol. 11, no. 6, p. 299, 2023, doi: 10.11114/smc.v11i6.6214.

[3] R. S. Daud, “The Role of Political Communication in Shaping Public Opinion: A Comparative Analysis of Traditional and Digital Media,” *J. Public Represent. Soc. Provis.*, vol. 1, no. 2, pp. 63–69, 2021, doi: 10.55885/jprsp.v1i2.241.

[4] M. Adamik-Szysiak, “Social Media as a Tool of Political Permanent Campaign on the Example of the Activity of Polish Politicians,” *SWS J. Soc. Sci. Art*, vol. 1, no. 2, pp. 1–12, 2019, doi: 10.35603/ssa2019/issue2.01.

[5] R. Karlsen and B. Enjolras, “Styles of Social Media Campaigning and Influence in a Hybrid Political Communication System,” *Int. J. Press.*, vol. 21, no. 3, pp. 338–357, 2016, doi: 10.1177/1940161216645335.

[6] A. Afandi, “Implications of the Digital Age on Social Engagement and Political Participation in the Context of Civic Education,” *J. Pendidik. Amartha*, vol. 2, no. 2, pp. 112–120, 2023, doi: 10.57235/jpa.v2i2.666.

[7] M. Rahbarqazi and S. Nourbakhsh, “The Effect of Social Media on Iranian Citizens’ Electoral Participation and Political Action,” *Commun. Soc.*, pp. 89–95, 2023, doi: 10.15581/003.36.2.83-95.

[8] A. Saleh, “Do Youth Trust in Political News on Social Media? Impacts of Trust or Distrust on the Egyptian Young Citizens Political Engagement,” □□□□□□

[illegible], vol. 19, no. 1, pp. 47–86, 2020, doi: 10.21608/joa.2020.127814.
[9] C. B. Tenorio, "Viral Voices: Exploring Twitter as a Platform for Public Engagement in the 2022 Philippine Election," *J. Gov. Polit.*, pp. 305–325, 2023, doi: 10.18196/jsp.v14i3.341.
[10] R. Gerodimos and J. Justinussen, "Obama's 2012 Facebook Campaign: Political Communication in the Age of the Like Button," *J. Inf. Technol. Polit.*, vol. 12, no. 2, pp. 113–132, 2014, doi: 10.1080/19331681.2014.982266.
[11] V. M. Rumata and F. Nugraha, "An Analysis of Fake Narratives on Social Media During 2019 Indonesian Presidential Election," *J. Komun. Malays. J. Commun.*, vol. 36, no. 4, pp. 351–368, 2020, doi: 10.17576/jkmjc-2020-3604-22.
[12] A. A. Firdaus, A. Yudhana, and I. Riadi, "Public Opinion Analysis of Presidential Candidate Using Naïve Bayes Method," *Kinet. Game Technol. Inf. Syst. Comput. Netw. Comput. Electron. Control*, 2023, doi: 10.22219/kinetik.v8i2.1686.
[13] S. F. Pratama and A. M. Wahid, "Fraudulent Transaction Detection in Online Systems Using Random Forest and Gradient Boosting," *J. Cyber Law*, vol. 1, no. 1, Art. no. 1, Mar. 2025.
[14] D. Arifka, M. N. Hakim, A. S. Adhipta, K. S. S. Yogananda, R. Salsabila, and R. Ferdiana, "Pandemic Fatigue: An Analysis of Twitter Users' Sentiments Against the COVID-19 in Indonesia," *J. Psikol.*, vol. 49, no. 2, p. 182, 2022, doi: 10.22146/jpsi.71979.
[15] J. Xue *et al.*, "Twitter Discussions and Emotions About the COVID-19 Pandemic: Machine Learning Approach," *J. Med. Internet Res.*, vol. 22, no. 11, p. e20550, 2020, doi: 10.2196/20550.
[16] S. Barreto, R. Moura, J. Carvalho, A. Paes, and A. Plastino, "Sentiment Analysis in Tweets: An Assessment Study From Classical to Modern Word Representation Models," *Data Min. Knowl. Discov.*, vol. 37, no. 1, pp. 318–380, 2022, doi: 10.1007/s10618-022-00853-0.
[17] Q. Alvi, "On the Frontiers of Twitter Data and Sentiment Analysis in Election Prediction: A Review," *Peerj Comput. Sci.*, vol. 9, p. e1517, 2023, doi: 10.7717/peerj-cs.1517.
[18] M. B. Alvi, A. Naeem, A. Mukhtiar, and M. H. Akram, "An Effective Framework for Tweet Level Sentiment Classification Using Recursive Text Pre-Processing Approach," *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, no. 6, 2019, doi: 10.14569/ijacsa.2019.0100674.
[19] K. Ayyub, S. Iqbal, E. U. Munir, M. W. Nisar, and M. Abbasi, "Exploring Diverse Features for Sentiment Quantification Using Machine Learning Algorithms," *Ieee Access*, vol. 8, pp. 142819–142831, 2020, doi: 10.1109/access.2020.3011202.
[20] M. Ahmad, S. Aftab, and I. Ali, "Sentiment Analysis of Tweets Using SVM," *Int. J. Comput. Appl.*, vol. 177, no. 5, pp. 25–29, 2017, doi: 10.5120/ijca2017915758.
[21] E. a. S. Roja, "Performance of Machine Learning Models in Predicting Sentiments of Post-Covid Patients," *Int. J. Recent Innov. Trends Comput. Commun.*, vol. 11, no. 10, pp. 2324–2329, 2023, doi: 10.17762/ijritcc.v11i10.8953.
[22] P. Adititya, U. Enri, and I. Maulana, "Analisis Sentimen Ulasan Pengguna Aplikasi Myim3 Pada Situs Google Play Menggunakan Support Vector Machine," *Jurikom J. Ris. Komput.*, vol. 9, no. 4, p. 1020, 2022, doi: 10.30865/jurikom.v9i4.4673.
[23] S. H. Imanuddin, "Sentiment Analysis on Satusehat Application Using Support Vector Machine Method," *J. Electron. Electromed. Eng. Med. Inform.*, vol. 5, no. 3, pp. 143–149, 2023, doi: 10.35882/jeemi.v5i3.304.
[24] I. Fitriani, "Tweeting the Economy: Analyzing Social Media Sentiments and Macroeconomic Indicators," *J. Innov. Bus. Econ.*, vol. 7, no. 02, pp. 129–138, 2023, doi: 10.22219/jibe.v7i02.28227.
[25] F. G. Conrad, J. A. Gagnon-Bartsch, R. A. Ferg, M. F. Schober, J. Pasek, and E. Hou, "Social Media as an Alternative to Surveys of Opinions About the Economy," *Soc. Sci. Comput. Rev.*, vol. 39, no. 4, pp. 489–508, 2019, doi: 10.1177/0894439319875692.
[26] C. E. W. Utomo, M. Hariadi, and S. Sumpeno, "Clustering Data National

Examinations Based on Social Media Using K-Means Methods," *Jaree J. Adv. Res. Electr. Eng.*, vol. 4, no. 2, 2020, doi: 10.12962/j25796216.v4.i2.152.