ISSN 2722-2039

Sentiment Analysis of Instagram User Comments related to the Inauguration of Mr. Prabowo Subianto as President of the Republic of Indonesia Using Natural Language Processing

Jose Julian Hidayat [a,1,*], Cindy Setyowati [a], Aditya Pratama Werdana [a]

[a] Informatics Engineering, Pelita Bangsa University, Bekasi, Indonesia
[1] josejulianhidayat@gmail.com
* corresponding author

ARTICLE INFO

Article history
Received March 18, 2025
Revised September 11, 2025
Accepted November 23, 2025

Keywords
Sentiment analysis
Natural language processing
Logistic regression
Indonesian text mining
Political sentiment analysis
Social media sentiment
Text preprocessing

ABSTRACT

This study employs Natural Language Processing (NLP) techniques to analyze sentiments expressed in Instagram comments about Prabowo Subianto's inauguration as Indonesia's president. The dataset comprises a rich collection of user-generated comments, meticulously preprocessed with the *Sastrawi* stemmer tailored for Indonesian. This preprocessing stage includes rigorous text cleaning, stemming, and stopword removal, ensuring that the analysis is based on the most relevant linguistic elements. To accurately classify the sentiment of these comments as positive or negative, a logistic regression model has been trained. The model leverages TF-IDF (Term Frequency-Inverse Document Frequency) for effective feature extraction, enhancing the precision of the analysis. With promising results, particularly in identifying uplifting remarks that celebrate the new president's ascendance, this study underscores the essential role of natural language processing in unraveling public sentiment surrounding pivotal political events. The findings of this research not only shed light on the intricate tapestry of public opinion but also pave the way for future sentiment analysis endeavors within the vibrant landscape of Indonesian social media. The model demonstrates robust accuracy, illustrating its effectiveness in interpreting the nuanced sentiments of digital discourse surrounding significant political milestones.

1. Introduction

Social media has emerged as the primary platform for individuals to voice their thoughts about social and political events in recent years [1]. Real-time public opinion monitoring is possible because platforms like Instagram provide direct public participation through comments. To extract emotional information from comment data, such as responses to Mr. Prabowo Subianto's inauguration as President of the Republic of Indonesia, sentiment analysis is crucial [2]. Making decisions and creating policies can be aided by this knowledge of public opinion. Large volumes of text data can now be analyzed automatically and effectively due to advances in Natural Language Processing (NLP) [3]. NLP makes it straightforward to classify sentiment as positive, negative, or neutral, which is useful given the daily volume of thousands of Instagram user comments [4]. One well-known feature extraction method for extracting significant terms from unstructured data is TF-IDF (Term

Frequency-Inverse Document Frequency) [5]. The goal of this study is to precisely identify public mood patterns using the TF-IDF and Support Vector Machine (SVM) algorithms.

When applied to politics, sentiment analysis offers a profound understanding of how the public views public leaders and governmental initiatives [6]. According to earlier studies, political events such as elections and formal inaugurations frequently elicit diverse viewpoints [7]. The public uses social media to criticize political figures or policies and to express support for them [8]. Thus, examining the mood of Prabowo Subianto's inauguration might reveal how Indonesians are adjusting to the new government. However, there are particular difficulties with sentiment analysis on social media. It might be challenging to identify sarcasm or innuendo in certain statements using simple NLP methods alone [9]. Furthermore, social media data are often imbalanced, with a disproportionate number of positive or negative remarks, which may compromise the model's accuracy [10]. To maximize sentiment classification performance, this work employs oversampling and model fine-tuning.

This study also demonstrates the significance of social media as a political instrument. According to earlier research, social media platforms such as Instagram and Twitter are increasingly used to shape public perception and influence public opinion [11]. Therefore, the findings of this study are expected not only to shed light on public sentiment regarding Prabowo's inauguration but also to guide political actors in developing more effective communication campaigns [12]. The overall goal of this study is to use a mix of natural language processing (NLP) and optimum classification approaches to analyze public responses to Prabowo Subianto's inauguration. This study uses Instagram data to address the fundamental question of how the public views the new leadership and whether this perception supports the legitimacy of the new administration [13].

2. Materials and Methods

2.1 Natural Language Processing

In artificial intelligence, the study of how computers and human language interact is known as natural language processing, or NLP. NLP employs a range of methods, including tokenization, stemming, lemmatization, and parsing, to enable machines to interpret, comprehend, and produce human text or speech more effectively [40]. NLP is crucial to contemporary applications, including sentiment analysis, chatbots, and machine translation. NLP has been transformed by transformer models such as BERT and GPT, which provide a richer understanding of context. Transformers enable models to account for phrase structure, leading to improved performance on tasks such as text summarization and named entity recognition (NER) [41]. Models like Word2Vec and GloVe, which represent words as continuous vectors in a multidimensional space, are also used in contextual embedding-based natural language processing [42].

NLP is increasingly used across a variety of industries, including law, public services, and health. NLP is used in public services to assist the government in comprehending social media complaints and comments [43]. NLP is used in the healthcare industry to process medical data and identify trends that inform illness forecasting. NLP is also used in legal research to analyze legal documents and identify significant rulings. Automatic language translation is another area in which NLP is crucial. Instead of translating a phrase word by word, systems such as Google Translate and DeepL employ transformer-based models to translate text by considering the entire sentence context [40][41]. This makes translation more natural and precise, particularly for languages such as Arabic and Mandarin, which have intricate structures. The relevance of NLP to cross-cultural communication is strengthened by the ongoing refinement of transformer-based models to better account for cultural variation and idiomatic meanings [42].

Handling ambiguous material, which is frequently difficult in natural language analysis, is another area in which NLP is crucial. When a single word or phrase has many meanings depending on the context, ambiguity results. To address this, transformer-based models such as BERT and GPT have been trained to account for a sentence's broader context, enabling them to distinguish the meaning of the same word across contexts [40]. For instance, depending on context, the term "bank" may refer to either a financial institution or a riverbank. This feature enables NLP models to provide more accurate interpretations of ambiguous text, thereby increasing the precision of tasks such as sentiment analysis and machine translation [41], [42]. NLP still faces significant obstacles, primarily in processing ambiguous language and interpreting context. Explainable AI (XAI), which enables users to understand how models are produced, remains an area of active research [38]. To maintain the impartiality and reliability of AI-based systems, NLP also requires robust validation methods and high-quality datasets.

2.2 Data Collection

The IG Comments Export Tool, which enables the automated extraction of data from user comments on Instagram, was used to collect the data. This method has been used extensively in studies that analyze real-time social media data, primarily to track public sentiment regarding social and political events [14]. This method increases the efficiency and precision of data collection by providing rapid access to hundreds of comments without requiring manual entry [15]. Identifying pertinent data is one of the major obstacles in data collection. Because Instagram is primarily an image-based site, it frequently contains hashtags, emoticons, and unstructured text that must be filtered throughout the scraping process [16]. As a result, the first step is to filter the data, eliminating entries that aren't relevant to the inauguration, including spam or remarks [17].

Because the inaugural ceremony is ceremonial, the data frequently reveals an imbalance in mood, with positive remarks predominating. Because of this, sentiment analysis is prone to bias, which is why methods such as oversampling are necessary to balance sentiment-class distributions [18]. Oversampling helps the model identify patterns in less frequent but more significant negative comments for examination [19]. Furthermore, the Instagram dataset includes punctuation, emoticons, and URLs, all of which must be removed to improve the effectiveness of NLP models [20]. Despite being a crucial component of communication, emojis are frequently removed during preprocessing because conventional text-based models, such as TF-IDF and SVM, struggle to handle them [21].

The dataset is structured for feature extraction and subsequent analysis following the cleaning step. To ensure the model can be appropriately assessed, the dataset is then split into training and test sets. The cleansed Instagram data can provide precise insights into public reactions to the presidential inauguration, owing to this meticulous preparation [22]. See Table 1 below, which summarizes data from the dataset on Instagram user comments about one of Pak Prabowo's posts on the day the 2024–2029 president-elect was announced.

Table 1. Dataset Collection

No	Comment_Id	Username	Text
1	18047633098999587	megaseptiautami	Semoga makmur dan apa yg kita semua harapkan akan terwujud yg pnting BP sehat panjang umur. semoga bisa bertemu bapa ..salam Karawang jawabarat
2	17855957781242122	gitadlr_	Seluruh rakyat indonesia menaruh harapan besar kepada Bapak. Selamat bertugas pak. Sehat selalu
3	17981750051766550	lliiyyaa970	Akhirnyaaa selamat pak semoga amanah dalam menjalankan tugas
4	17977663406777314	ernasriyanu	Selamat bertugas bapak...semoga Amanah,kami percaya bapak bisa membawa Indonesia jauuuh lebih baik..jadilah suri tauladan yg baik..kelak saya bersaksi di akhirat bapak adalah pemimpin yang Amanah.. Aamiin..ɪᴅ
5	18461818831043515	masdomar	Selamat bertugas pak @prabowo, semoga selalu dalam bimbingan dan petunjuk Allah SWT.
…	…	…	…
…	…	…	…
9715	17994161273708905	fardan_ardiantoo	Semangat pak
9716	18063008275680935	kenang_af	Selamat bekerja pak
9717	17858743839269916	citrayulliw	Selamat bpk prabowo

9718	18030860474461282	kakzallls	Semangat bapak
9719	18061435924700448	wpryd	Presiden Ku
9720	18032183993076434	muhammadaqilnf	Selamat Bertugas Pak

2.3 Data Preprocessing

In sentiment analysis, preprocessing is an essential step, particularly when dealing with unstructured data like Instagram comments [14]. The objective of this step is to prepare the raw text for analysis by NLP models. Data cleaning is the initial stage of preprocessing, which eliminates superfluous components, including URLs, punctuation, and symbols [15]. This method ensures that only significant information is included in the analysis and helps reduce data noise. Furthermore, because emojis are often unintelligible to conventional text-based NLP algorithms such as TF-IDF, eliminating them is a crucial first step. According to some research, removing emojis can improve model performance by reducing ambiguity in the data, even when they carry specific emotional connotations [16]. Although some methods convert emojis into descriptive words, this approach is rarely used due to the additional complexity it introduces [17].

The words are then reduced to their base form using stemming and lemmatization. Stemming reduces variants such as "walk," "berjalanlah," and "berjalan-jalan" to "jalan" using the Sastrawi library to restore Indonesian words to their base form [18]. In addition to ensuring that terms with similar meanings are not handled as distinct entities, this step aids in data consistency [19]. Common words such as "and," "or," and "which" that don't contribute much to sentiment analysis can also be removed using stopword removal [20]. By focusing on more relevant terms, research indicates that eliminating stop words can improve model performance [21]. This is particularly crucial when working with comment data that frequently uses common idioms and conjunctions.

In preprocessing, case folding is employed in addition to stopword removal and stemming. To maintain consistency throughout the analysis, case folding converts the entire text to lowercase [23]. This is crucial because words with identical meanings may be interpreted differently when they are not standardized in format, and capitalization is often inconsistent in Bahasa Indonesia and on social media [24]. Case folding prevents misunderstanding during feature extraction and sentiment classification by treating words like "Selamat" and "selamat" as the same.

Tokenization, the last step, divides the text into discrete word units (tokens) for additional analysis [22]. This preprocessing produces a clean, consistent dataset suitable for TF-IDF feature extraction and subsequent analysis using algorithms such as SVM or Logistic Regression [2]. The model may generate more precise and pertinent predictions in determining public mood with appropriate preprocessing [4]. The data validation procedure is also crucial in preprocessing to guarantee that the data is prepared for use. The purpose of validation is to determine whether the data meet specific requirements, such as minimum text length and topical relevance [25]. Since they don't provide valuable emotional information, brief comments or emoticons without text are mostly ignored [26]. By ensuring that the model is trained only on pertinent data, this validation helps improve data quality and enhances the model's performance in complex sentiment analysis.

2.4 Sentiment Labeling

The practice of categorizing text into sentiment groups, such as neutral, negative, or positive, is known as sentiment labeling. This method is frequently used in Indonesian public opinion research on social and political issues, including elections and government regulations [27]. Researchers may use unstructured data from social media platforms like Instagram and Twitter to gain insight into people's reactions in real time [28]. The sentiment labeling technique relies extensively on both automated and semi-automatic models. The best options are Support Vector Machine (SVM), Naïve Bayes, and Logistic Regression because of their effectiveness and capacity to process introductory structured text [29]. Nevertheless, models based on IndoBERT, which were developed specifically for Bahasa Indonesia, are better at capturing context and mitigating ambiguity in social data, which frequently contains sarcasm [30].

Research conducted in Indonesia indicates that IndoBERT is highly proficient in sophisticated text analysis, including political data such as public opinion on presidential candidates and election delays [22]. Moreover, oversampling approaches can be used to address imbalanced datasets, thereby improving the accuracy of predicting minority attitudes, including negative feelings [30]. This method reduces prejudice arising from the unequal distribution of sentiment groups.

Small datasets are frequently manually labeled to ensure the high quality of the training data before use in automated methods [28]. Semi-automated labeling is used in specific research to expedite the categorization process while preserving high accuracy. This guarantees accurate, timely, and locally contextualized sentiment analysis [22]. Sentiment labeling is challenging in Indonesia due to data imbalance and variation in informal language. However, sentiment analysis can provide comprehensive insights for stakeholders, particularly in politics, when preprocessing methods, transform-based models, and data balancing procedures are employed [29].

2.5 Feature Extraction using TF-IDF

A popular feature extraction method in text analysis, particularly in emotion analysis and opinion categorization, is called TF-IDF (Term Frequency-Inverse Document Frequency). According to a study conducted in Indonesia, TF-IDF and Support Vector Machine (SVM) jointly improve the categorization accuracy of e-commerce reviews, including Shopee product reviews [31]. This illustrates how well TF-IDF filters out standard, meaningless terms and captures keywords pertinent to user opinions. Public opinion regarding government policies and responses to social issues are also analyzed using TF-IDF in the context of public policy and analysis. For instance, studies on the COVID-19 epidemic demonstrate that TF-IDF facilitates the extraction of themes related to public mood from social media [32]. This shows the applicability of the TF-IDF approach to managing unstructured content generated in real time on social media platforms such as Instagram and Twitter.

TF-IDF is used in false news identification analysis in the United States. In contrast to word-embedding-based techniques, research has shown that combining TF-IDF with logistic regression yields strong performance in detecting false news [33]. This technique enables more precise categorization and aids in managing news items with comparable content but minor variations in key phrases. According to studies conducted in Europe, TF-IDF is frequently used for political sentiment analysis and stock market analysis. Essential elements from financial news were extracted using TF-IDF in research on how the stock market reacts to economic news [34]. As a result, by using this method, analysts and investors may identify attitudes that could influence stock prices and use news trends to inform more informed judgments. TF-IDF has drawbacks despite its effectiveness and adaptability across contexts. The incapacity of TF-IDF to capture contextual meaning or word relationships is one of its drawbacks [30]. To handle data with more complex contexts and richer semantics, several studies in the United States and Europe have integrated TF-IDF with transformer models such as BERT.

2.6 Model Training, Testing, and Evaluation

For the model to provide precise predictions, it must first identify patterns in the training data. To ensure the model doesn't overfit to specific data, cross-validation is widely used in America. This method splits the dataset into folds, with each fold serving as both training and test data [35]. Cross-validation reduces the model's potential bias and yields more generalizable findings [36]. Following training, a previously untested dataset is used to test the model. To objectively evaluate model performance, evaluation criteria such as accuracy, precision, recall, and F1-score are frequently employed. Since the F1-score balances accuracy and recall, it is beneficial for unbalanced datasets [37]. This evaluation confirms the model's ability to perform successfully on new data, rather than merely recalling the training data.

Model interpretability is a top priority in America, particularly in sensitive fields such as banking and health care. In order to provide transparency and guarantee that users comprehend the choices made by AI, models like SHAP (SHapley Additive exPlanations) are utilized to describe the aspects that impact the prediction outcomes [38]. Because they ensure ethical deployment of automated systems, explainable AI (XAI) models are becoming increasingly common [35]. The model's resilience to shifts in data patterns, or "concept drift," is examined in addition to its initial testing. These shifts are typical in social and financial data research, as market trends or people's behavior might shift [36]. Accuracy will be lost by models that do not adjust rapidly. Consequently, to sustain performance, retraining with fresh data is required regularly [39]. Before deploying models in the real world, several American businesses, including the military and aviation, employ operational simulation. Before the system is extensively used, these simulations enable developers to identify any problems and make necessary corrections [37]. This method ensures that the model is reliable across a range of operating conditions and is ready for use in real-world settings.

2.7 Deployment of Model

A sentiment analysis model was developed for this project, using data scraped from Instagram comments posted during Mr. Prabowo Subianto's inauguration. When data is not accessible through public APIs, scraping techniques enable rapid and effective data collection [44]. This deployment methodology uses batches of scraped data, which are analyzed regularly rather than immediately integrated into social media platforms. Since data may be gathered, cleansed, and analyzed after an incident, this approach is pertinent to post-event analysis. Deploying sentiment models requires constant maintenance and observation. Frequent retraining of the model is necessary to detect and correct concept drift or shifts in sentiment patterns over time. To ensure that model performance remains optimal, cloud platforms provide capabilities for automatic monitoring, anomaly detection, and alerting operational personnel [45]. The model can adapt to shifting social dynamics and still yield accurate findings with proper tracking.

3. Results and Discussion

The model's performance is reported in the Classification Report using several key metrics, including accuracy, recall, and the F1 Score for each class (positive and negative). The evaluation's findings were as follows:

Table 2. Findings' Evaluation

Metric	Negative Sentiment (0)	Positive Sentiment (1)	Weighted Avg
Precision	0.58	0.91	0.88
Recall	0.04	1.00	0.91
F1-Score	0.07	0.95	0.87
Accuracy			0.91

Table 3. Confusion Matrix

	Predicted Negative (0)	Predicted Positive (1)
Actual Negative (0)	7	173
Actual Positive (1)	5	1759

Evaluates the accuracy of optimistic forecasts. In other words, how many of the optimistic forecasts came true?

1. Accuracy

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$= \frac{1759 + 7}{1759 + 7 + 173 + 5} = \frac{1766}{1944} = 0{,}91(91\%)$$

This indicates that the model's classification accuracy for the overall mood of comments is 91%.

2. Precision

$$Precision = \frac{TP}{TP + FP} = \frac{1759}{1759 + 173} = \frac{1759}{1932} = 0{,}91$$

Precision gauges the model's accuracy in forecasting positive sentiment. Therefore, 91% of all optimistic forecasts turned out to be positive.

3. Recall

$$Recall = \frac{TP}{TP + FN} = \frac{1759}{1759 + 5} = \frac{1759}{1764} = 1{,}00$$

The model's recall gauges its ability to identify every genuine positive instance. It demonstrates the model's ability to identify all favorable comments.

4. F1-Score

The F1-score is the accuracy and recall of harmonic mean.

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

$$F1 = 2 \times \frac{0{,}91 \times 1{,}00}{0{,}91 + 1{,}00} = 2 \times \frac{0{,}91}{1{,}91} = 0{,}95$$

With an F1-score of 95%, the model shows a reasonable balance between recall and accuracy.

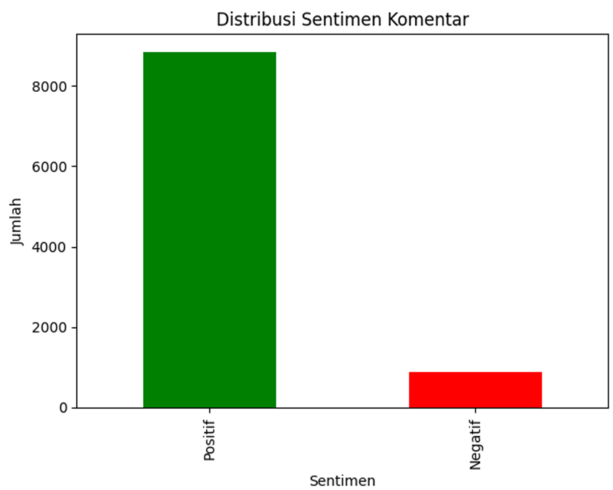

Figure 1. Comment sentiment distribution

Acknowledgment

We thank all the individuals and organizations that helped make this study possible. We would especially like to thank the creators of the IG Comments Export tool for making it possible to extract data for our investigation. We also thank Scikit-learn, NLTK, *Sastrawi*, and Matplotlib, the creators of the Python libraries used in this study, for their tools and documentation, which enabled this analysis. Lastly, we would like to thank Pelita Bangsa University for its ongoing assistance in providing the facilities and resources required to finish this study.

References

[1] K. L. Tan *et al.*, "A survey of sentiment analysis: Approaches, datasets, and future research," *Appl. Sci.*, vol. 13, no. 7, p. 4550, Apr. 2023, doi: 10.3390/app13074550.

[2] H. A. Alatabi and A. R. Abbas, "Sentiment analysis in social media using machine learning techniques," *Iraqi J. Sci.*, vol. 61, no. 1, pp. 193–201, Jan. 2020, doi: 10.24996/ijs.2020.61.1.22.

[3] A. Kampatzis, A. Sidiropoulos, K. Diamantaras, and S. Ougiaroglou, "Sentiment dimensions and intentions in scientific analysis: Multilevel classification in text and citations," *Electronics*, vol. 13, no. 9, p. 1753, May 2024, doi: 10.3390/electronics13091753.

[4] Z. Zhou, J. Qin, X. Xiang, Y. Tan, Q. Liu, and N. N. Xiong, "News text topic clustering optimized method based on TF-IDF algorithm on Spark," *Comput., Mater. Contin.*, vol. 62, no. 1, pp. 217–231, 2020, doi: 10.32604/cmc.2020.06431.

[5] P. Guleria, J. Frnda, and P. N. Srinivasu, "NLP based text classification using TF-IDF enabled fine-tuned long short-term memory: An empirical analysis," *Array*, vol. 27, p. 100467, Sep. 2025, doi: 10.1016/j.array.2025.100467.

[6] A. Albladi, M. Islam, and C. Seals, "Sentiment analysis of Twitter data using NLP models: A comprehensive review," *IEEE Access*, vol. 13, pp. 30444–30468, 2025, doi: 10.1109/ACCESS.2025.3541494.

[7] A. A. Ghafoor and M. A. Aish, "Enhanced fake news detection with Bi-LSTM networks and TF-IDF," *J. Innov. Comput. Emerg. Technol.*, vol. 4, no. 2, pp. 1–10, Oct. 2024, doi: 10.56536/jicet.v4i2.124.

[8] M. Arman *et al.*, "Evaluating transformer-based models for Bengali political sentiment analysis," in *Proc. 5th Int. Conf. Electr., Comput. Energy Technol. (ICECET)*, Paris, France, Jul. 2025, pp. 1–6. [Online]. Available: https://hal.science/hal-05068202

[9] X. Qi and A. Shabrina, "Twitter sentiment analysis using NLP," *Ain Shams Eng. J.*, vol. 14, no. 9, p. 102112, Sep. 2023, doi: 10.1049/icp.2023.2562.

[10] M. Trevisan *et al.*, "Towards understanding political interactions on Instagram," in *Proc. 30th ACM Conf. Hypertext Soc. Media (HT)*, Hof, Germany, Sep. 2019, pp. 247–251, doi: 10.1145/3342220.3343657.

[11] W. N. Chan and T. Thein, "A comparative study of machine learning techniques for real-time multi-tier sentiment analysis," in *Proc. 1st IEEE Int. Conf. Knowl. Innov. Invention (ICKII)*, Jeju, Korea (South), 2018, pp. 90–93, doi: 10.1109/ICKII.2018.8569169.

[12] H. Bashiri and H. Naderi, "Comprehensive review and comparative analysis of transformer models in sentiment analysis," *Knowl. Inf. Syst.*, vol. 66, no. 12, pp. 7305–7361, Dec. 2024, doi: 10.1007/s10115-024-02214-3.

[13] K. Srinidhi, A. Harika, S. S. Voodarla, and J. Abhiram, "Sentiment analysis of social media comments using natural language processing," in *Proc. Int. Conf. Inventive Comput. Technol. (ICICT)*, Lalitpur, Nepal, 2024, pp. 762–768, doi: 10.1109/ICICT60155.2024.10544849.

[14] S. Pappala, "Sentiment-driven movie recommendation system: A machine learning approach," *Int. J. Multidiscipl. Res. (IJFMR)*, vol. 6, no. 3, pp. 1–6, May-Jun. 2024.

[15] S. Gnanapriya and K. Anandan, "Sentiment analysis in political discourse: Understanding public opinion from social media," *E-Learn. Digit. Media*, Early Access, 2025, doi: 10.1177/20427530251343205.

[16] S. Pradha, M. N. Halgamuge, and N. Tran Quoc Vinh, "Effective text data preprocessing technique for sentiment analysis in social media data," in *Proc. 11th Int. Conf. Knowl. Syst. Eng. (KSE)*, Da Nang, Vietnam, 2019, pp. 1–8, doi: 10.1109/KSE.2019.8919368.

[17] K. Al Sharou, Z. Li, and L. Specia, "Towards a better understanding of noise in natural language processing," in *Proc. Int. Conf. Recent Adv. Nat. Lang. Process. (RANLP)*, Held Online, Sep. 2021, pp. 53–62. [Online]. Available: https://aclanthology.org/2021.ranlp-1.7/

[18] H. T. Duong and T. A. Nguyen-Thi, "A review: Preprocessing techniques and data augmentation for sentiment analysis," *Comput. Soc. Netw.*, vol. 8, no. 1, p. 1, Jan. 2021, doi: 10.1186/s40649-020-00080-x.

[19] Q. A. Xu, V. Chang, and C. Jayne, "A systematic review of social media-based sentiment analysis: Emerging trends and challenges," *Decis. Analytics J.*, vol. 3, p. 100073, Jun. 2022, doi: 10.1016/j.dajour.2022.100073.

[20] W. Xie, W. Zhao, and Z. Xu, "Comparative analysis of multi-platform e-commerce online reviews based LDA and sentiment dictionary," *Int. J. Intell. Comput. Cybern.*, vol. 18, no. 2, pp. 298–325, 2025, doi: 10.1108/IJICC-08-2024-0384.

[21] M. Alfreihat *et al.*, "Emo-SL framework: Emoji sentiment lexicon using text-based features and machine learning for sentiment analysis," *IEEE Access*, vol. 12, pp. 81793–81812, 2024, doi: 10.1109/ACCESS.2024.3382836.

[22] S. Dorle and N. Pise, "Political sentiment analysis through social media," in *Proc. 2nd Int. Conf. Comput. Methodol. Commun. (ICCMC)*, Erode, India, 2018, pp. 869–873, doi: 10.1109/ICCMC.2018.8487879.

[23] E. S. Tellez, S. Miranda-Jiménez, M. Graff, D. Moctezuma, O. S. Siordia, and E. A. Villaseñor, "A case study of Spanish text transformations for twitter sentiment analysis," *Expert Syst. Appl.*, vol. 81, pp. 457–471, Sep. 2017, doi: 10.1016/j.eswa.2017.03.071.

[24] S. R. El-Beltagy and A. Ali, "Open issues in the sentiment analysis of Arabic social media: A case study," in *Proc. 9th Int. Conf. Innov. Inf. Technol. (IIT)*, Al Ain, United Arab Emirates, 2013, pp. 215–220, doi:10.1109/Innovations.2013.6544421.

[25] J. Lappeman, R. Clark, J. Evans, L. Sierra-Rubia, and P. Gordon, "Studying social media sentiment using human validated analysis," *MethodsX*, vol. 7, p. 100867, 2020, doi: 10.1016/j.mex.2020.100867.

[26] U. Naseem, I. Razzak, and P. W. Eklund, "A survey of pre-processing techniques to improve short-text quality: A case study on hate speech detection on Twitter," *Multimedia Tools Appl.*, vol. 80, no. 28, pp. 35239–35266, Nov. 2021, doi: 10.1007/s11042-020-10082-6.

[27] A. Perdana *et al.*, "Analisis sentimen terhadap isu penundaan pemilu di Twitter menggunakan Naive Bayes classifier," *Jurnal Sisfokom*, vol. 11, no. 2, pp. 195–200, Jul. 2022, doi: 10.32736/sisfokom.v11i2.1412.

[28] L. Geni *et al.*, "Sentiment analysis of tweets before the 2024 elections in Indonesia using IndoBERT," *J. Inf. Technol. Comput. Sci.*, vol. 9, no. 3, pp. 746–757, 2023, doi: 10.26555/jiteki.v9i3.26490.

[29] S. M. P. Tyas *et al.*, "A robustly optimized BERT using random oversampling for analyzing imbalanced stock news sentiment data," in *Proc. Int. Conf. Comput. Sci., Inf. Technol., Electr. Eng. (ICCoSITE)*, Jakarta, Indonesia, Mar. 2023, pp. 134–139, doi: 10.1109/ICCoSITE57641.2023.10127725.

[30] B. Wilie *et al.*, "IndoNLU: Benchmark and resources for evaluating Indonesian natural language understanding," in *Proc. 1st Conf. Asia-Pacific Chapter Assoc. Comput. Linguistics 10th Int. Joint Conf. Nat. Lang. Process. (AACL-IJCNLP)*, Suzhou, China, Dec. 2020, pp. 843–857, doi: 10.18653/v1/2020.aacl-main.85.

[31] E. H. Aydoğan and F. Y. Okay, "Sentiment analysis of reviews for E-commerce applications," in *Proc. 9th Int. Conf. Comput. Sci. Eng. (UBMK)*, Antalya, Turkiye, 2024, pp. 1–6, doi: 10.1109/UBMK63289.2024.10773400.

[32] N. Agustiana *et al.*, "Comparison of sentiment analysis on Traveloka and Tiket.com using Naïve Bayes method," *ITEJ (Inf. Technol. Eng. J.)*, vol. 8, no. 2, pp. 112–120, 2023, doi: 10.24235/itej.v8i2.119.

[33] V. V. V. R. Gurram and J. M. Nalam, "Automated detection of fake news in natural language processing: A comparative study of TF-IDF and lexical-based stance detection with logistic regression," Blekinge Inst. Technol., Karlskrona, Sweden, Tech. Rep., 2024, pp. 60.

[34] M. Tezgider *et al.*, "Bidirectional transformer models for financial sentiment analysis," *Concurrency Computat.: Pract. Exper.*, vol. 34, no. 20, p. e6486, Sep. 2022, doi: 10.1002/cpe.6486.

[35] Z. C. Lipton and J. Steinhardt, "Troubling trends in machine learning scholarship: Some ML papers suffer from flaws that could mislead the public and stymie future research," *Queue*, vol. 17, no. 1, pp. 45–77, Feb. 2019.

[36] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning with Applications in R*, 2nd ed. New York, NY, USA: Springer, 2021. doi: 10.1007/978-1-0716-1418-1.

[37] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.

[38] C. Molnar, *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*. 2018. [Online]. Available: https://christophm.github.io/interpretable-ml-book/

[39] J. D. Kelleher and B. Tierney, *Data Science*. Cambridge, MA, USA: MIT Press, 2018.

[40] J. H. Martin and D. Jurafsky, *Speech and Language Processing*. Upper Saddle River, NJ, USA: Prentice Hall, 2000.

[41] A. Vaswani *et al.*, "Attention is all you need," in *Proc. 31st Int. Conf. Neural Inf. Process. Syst. (NeurIPS)*, Long Beach, CA, USA, Dec. 2017

[42] IBM, "What is natural language processing (NLP)?," IBM, Armonk, NY, USA, Tech. Rep., 2024. [Online]. Available: https://www.ibm.com/topics/natural-language-processing.

[43] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," *arXiv:1301.3781*, 2013. [Online]. Available: https://arxiv.org/abs/1301.3781

[44] S. Mittal, A. Chandel, and P. M. Nguyen, "AI and social media analytics: Leveraging real-time data for entrepreneurial growth," in *Improving Entrepreneurial Processes Through Advanced AI*, IGI Global, 2025, pp. 267–290.

[45] M. Jindal, "Advancements in cloud-based machine learning: Navigating deployment and scalability," *Int. J. Multidiscip. Res.*, vol. 5, no. 6, 2023, doi: 10.36948/ijfmr.2023.v05i06.11371.