

Sentiment Analysis of Netizens' Comments About Dwi Sasetyaningsih on the TikTok Platform Using a Deep Learning-Based Method Long Short Term -Memory (LSTM)

Ditha Aditya Pernikasari *, Tristy Welas Ulas Jati

Department of Communication Studies, Padjadjaran University, 45363, Indonesia

*Corresponding author's email: fatiyaaza09@gmail.com

ABSTRACT

Keywords

Sentiment analysis; social media;
TikTok; LSTM; Bing Liu

The digital age, marked by rapid advancements in information technology, has transformed social media from merely a source of entertainment into a space where the public expresses opinions on various public issues. One such issue that has sparked a wide range of reactions on social media is the statement made by Dwi Sasetyaningsih, a recipient of the LPDP scholarship, regarding her child's citizenship. This study aims to analyze sentiment in text data by applying a deep learning method based on Long Short-Term Memory (LSTM), and drawing on Bing Liu's theory of Sentiment Analysis and Opinion Mining. The data used consists of a collection of comments from TikTok platform, which were then categorized into two sentiment classes: positive and negative. The research results produced an LSTM model from a dataset of 3,395 comments with a best accuracy of 87.48%. "positive" comments accounted for 72.90%, while "negative" comments accounted for 27.20%. It was concluded that the TikTok platform differs from other social media platforms in sentiment analysis of comment sections using text analysis methods.

1. Introduction

The digital age has transformed various aspects of society, particularly communication patterns, which have become increasingly interactive and media-driven (Tika et al., 2023). Social media is no longer merely a form of entertainment but has evolved into a participatory communication space that enables users to produce, distribute, and respond to information in real time (Knittel et al., 2021). This transformation reflects a paradigm shift in mass communication from a one-way model toward interactive communication based on user-generated content, in which the formation of public opinion is no longer centered on mass media institutions but develops dynamically through interactions among users (Hussain et al., 2023).

One of the social media platforms experiencing rapid growth is TikTok (Lestari, 2025).. This platform is characterized by interactive short-form video content and a comment section that allows users to respond directly and spontaneously. These characteristics make TikTok an active space for shaping and reflecting public opinion (Kencana et al., 2024). TikTok is also among the social media platforms with a large global user base (Gray, 2021).. Indonesia is among the countries with a high number of TikTok users, indicating a high level of user penetration and interaction with the platform

(Ariyus & Manongga, 2024). The high usage of TikTok makes this platform a strategic space for shaping public opinion on various emerging issues (Wei, 2025).

A phenomenon that has captured public attention emerged in a case involving Dwi Sasetyaningsih, a scholarship awardee from Lembaga Pengelola Dana Pendidikan (LPDP), who posted on social media about her child's citizenship. The post sparked a range of responses from netizens and spread across platforms, including TikTok, thereby contributing to the formation of public discourse in the digital space (Besman et al., 2018). This phenomenon demonstrates a process of information diffusion that reinforces the dynamics of public opinion within the social media ecosystem and reflects how personal issues can evolve into widespread public discourse (Wismashanti & Indonesia, 2023).

User responses on social media reveal a diverse range of public opinions that reflect differing attitudes toward a particular issue. Public opinion analysis helps identify how these responses form and evolve in the digital space (Naratif et al., 2023). One approach used to examine this phenomenon is sentiment analysis (Cas, 2023). Sentiment analysis categorizes users' opinions and emotions into specific categories, such as positive and negative. Bing Liu explains that opinions play a crucial role in influencing human behavior and decision-making, both in individual and organizational contexts (Liu, 2012).

With the rapid development of social media, the volume of opinions available online has increased significantly and is spread across various platforms, making manual analysis ineffective (Ahmad et al., 2025). Therefore, sentiment analysis is necessary to automatically and systematically extract opinions, emotions, and evaluations from textual data. This approach is used to identify and map sentiment patterns and reveal trends in public attitudes in response to issues emerging on social media, particularly the public's response to Dwi Sasetyaningsih regarding her child's citizenship status.

A number of previous studies in the field of sentiment analysis have generally focused on social media platforms such as Twitter, Instagram, and YouTube, using machine learning and deep learning methods (Liyih et al., 2024). However, most research remains dominated by data from Twitter and Instagram, while data from TikTok remain relatively limited (Zhang, 2023). Additionally, studies using TikTok data still largely rely on traditional classification methods such as Naïve Bayes (Nurjanah & Apidana, 2025) and have not yet extensively implemented specific deep learning approaches. On the other hand, sentiment analysis research on social media platforms largely focuses on general issues such as brand reputation, app reviews, and public comments, so there has not been much research specifically examining sentiment toward specific figures on social media (Abdi et al., 2019, (Analysis, 2024), (Wu & Hong, 2022), (Norhidayati, 2021).

Therefore, this study aims to analyze the sentiment of netizens' comments about Dwi Sasetyaningsih on the TikTok platform using a deep learning method based on Long Short-Term Memory (LSTM).

1.2 Sentiment Analysis in Digital Communication

Public opinion emerging on social media can be understood as a form of social interaction that develops within a participatory digital communication space, where individuals actively express their views and responses to an issue. To understand these dynamics, sentiment analysis is used to identify and categorize opinions and emotions in text data, thereby revealing trends in public responses. Bing Liu explains that public opinion in digital media influences how individuals make decisions. In the context of social media, user comments reflect public attitudes toward an issue (Liu, 2012).. The volume of text-based data on digital platforms, such as social media, continues to grow significantly, necessitating effective methods for systematic analysis. One widely used approach is deep learning, specifically Long Short-Term Memory (LSTM), which can analyze text data more accurately (Pusat et al., 2018). Therefore, this method is used in this study to analyze the sentiment of comments on the TikTok platform.

2. Method

This study employs a quantitative approach to analyze public sentiment in comments on the TikTok platform regarding the issue of Dwi Sasetyaningsih concerning her child's citizenship status. A deep learning-based sentiment analysis approach was used, employing a Long Short-Term Memory (LSTM) model to classify comments into positive and negative categories.

Long Short-Term Memory (LSTM) is a type of Recurrent Neural Network (RNN) capable of processing sequential data and retaining information over the long term. This capability enables LSTM to understand the context within text data, making it effective for sentiment analysis tasks (Mutmatinah, 2020).

Data collection in this study was conducted using automated web scraping techniques via the Apify Client. Data were obtained from user comments on news videos publicly posted on the TikTok platform. The initial dataset consisted of 3,842 comments. Subsequently, a data cleaning process was performed, which included removing duplicates, filtering out irrelevant comments, and removing comments containing URLs. After this process, the dataset used in this study comprised 3,395 comments. The research workflow for this study is presented in a flowchart, as shown in the following figure.

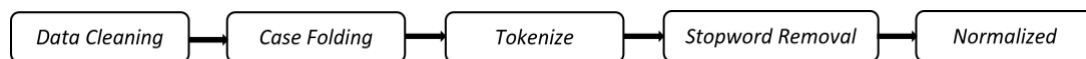


Figure 1. Research Steps

This research workflow is part of the data preprocessing stage. It includes several key steps: data cleaning, case folding, tokenization, stopwords removal, and normalization. The data cleaning stage involves removing irrelevant characters such as emojis, URLs, punctuation, numbers, and comments unrelated to the research topic. Next, case folding is performed by converting all text to lowercase to maintain data consistency. The normalization process converts non-standard words or slang into their standard forms. The text is then tokenized to break sentences into words, followed by stopwords removal using an Indonesian stopwords list to eliminate words that lack significant meaning.

3. Result and Discussion

3.1 Data Collection

The data for this study consists of TikTok user comments extracted from news videos publicly posted on the TikTok platform. The comment data collection process took place on April 7, 2026, and was conducted automatically using the Apify Client via web scraping, targeting video URLs with more than 2,000 comments. During the scraping process, a total of 3,842 initial data points were obtained; after cleaning, removing duplicate data and comments containing URLs or links, a dataset of 3,395 data points was obtained.

3.2 Data Preprocessing

To prepare comment text for data and computational analysis, preprocessing was performed. This process includes removing emojis, URLs, mentions, hashtags, numbers, and non-alphabetic characters using regular expressions, and converting the text to lowercase using case folding. It also converts slang and informal variations commonly used on TikTok into standardized words to normalize non-standard language. Next, the text is divided into tokens using tokenization techniques. Subsequently, an expanded lexicon with topic-specific stopwords is used to remove Indonesian stopwords.

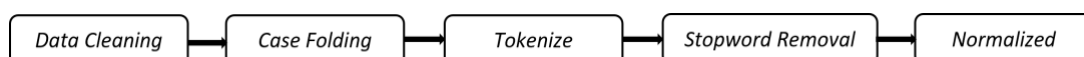


Figure 2. Data Preprocessing Steps

3.1 Data Analysis

Comment data is divided into two types: negative and positive sentiment. Positive sentiment indicates satisfaction or support, while negative sentiment includes criticism, complaints, and dissatisfaction. The process of classifying or labeling comments is carried out so that the LSTM model can accurately learn linguistic patterns and distinguish emotional polarity within text sequences.

Table 2. Labeling Result

<i>Comment</i>	<i>Cleaning</i>	<i>Case Folding</i>	<i>Tokenize</i>	<i>Stopword Removal</i>	<i>Normalized</i>
terlalu over sharing sii	terlalu over sharing sii	terlalu over sharing sii	[terlalu, over, sharing, sih]	over sharing sih	over sharing sih

Figure 2 presents the results of the sentiment analysis of TikTok comments, categorizing the data into two groups: positive and negative. Based on the visualization, the majority of comments are positive, with 2,475 data points (72.90%), while negative sentiment accounts for 920 data points (27.10%). This indicates that TikTok users' perceptions of the analyzed topic are predominantly positive.

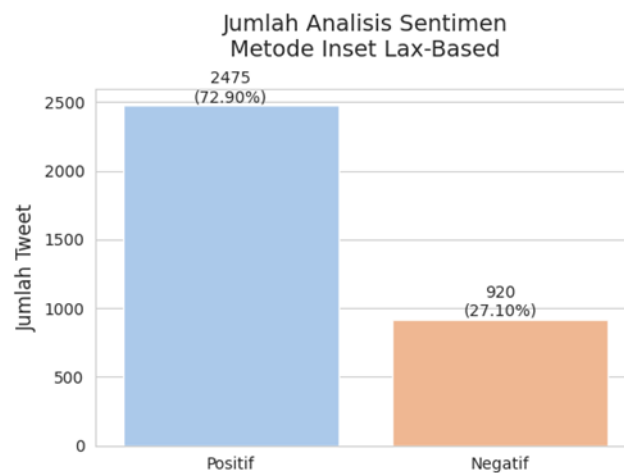


Figure 3. Percentage of negative and positive comments

The prevalence of positive comments may indicate several things. First, the content analyzed likely offers entertainment, value, or information that resonates with the audience. Second, TikTok's communication style, which tends to be lighthearted and expressive, may also encourage positive responses. Overall, this analysis suggests that public opinion on TikTok regarding this topic leans positive, although negative comments still exist.

Following the data preprocessing stage, the most frequent words are shown in the word cloud in Figure 3. Based on the sentiment analysis, the most dominant words are "indonesia," "beasiswa," "tidak," "uang," and "anak," indicating that public conversations on TikTok largely focus on scholarship issues and their connection to social conditions in Indonesia. The words "beasiswa" and "Indonesia" appear in the largest font sizes, signaling the main topics of discussion, while "tidak," which is also quite dominant, reflects a significant amount of negative sentiment or rejection in the comments. On the other hand, words such as "selamat," "terima," "bangga," and "cinta" indicate positive sentiment in the form of support and appreciation. Additionally, words like "uang," "pajak," and "pemerintah" indicate that many comments link the scholarship issue to the use of public funds, thereby sparking both support and opposition. Overall, this word cloud demonstrates that public opinion is diverse, with the dominant topic of scholarships in Indonesia triggering both positive sentiments and criticism from the public.



Figure 4. Wordcloud of frequently appearing words

The LSTM model achieved an accuracy of 87.48 percent, indicating that it is capable of accurately predicting sentiment. Overall, an accuracy of 87.48 percent can be considered a good result in sentiment analysis, especially when the data comes from social media platforms like TikTok, which tend to feature informal language, abbreviations, and a wide variety of expressions.

Based on the loss and accuracy plots, the LSTM model demonstrated better performance during training, as evidenced by a consistent decrease in training loss from approximately 2.0 to 0.35 and an increase in training accuracy to nearly 0.99, indicating that the LSTM model has a strong ability to learn patterns in the training data. On the validation data, the validation loss initially decreased until around the 6th or 7th epoch, but then rose again. The validation accuracy only increased to around 0.85 and tended to stagnate. This pattern indicates that overfitting occurs when the model adapts too closely to the training data, making it difficult to generalize to new data.

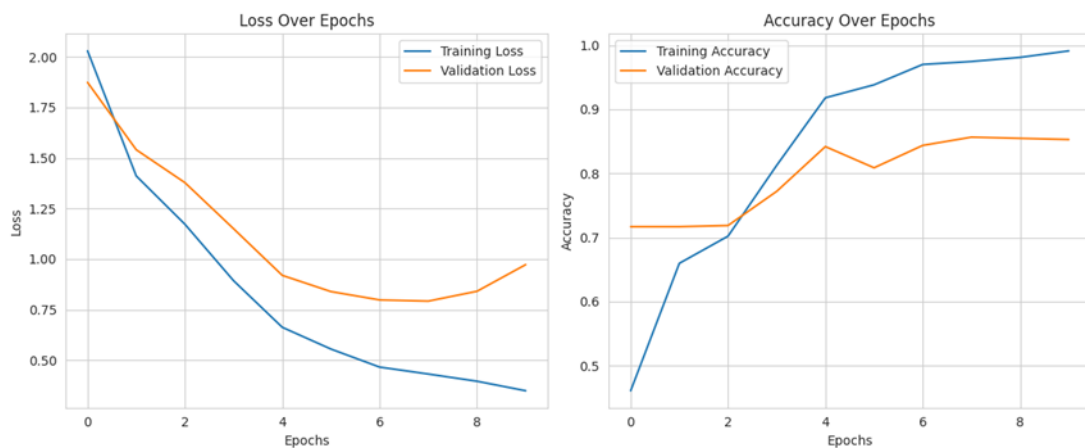


Figure 5. Actual and Prediction

4. Conclusion

Based on the analysis results, the LSTM model used in this study was able to classify sentiment effectively, as evidenced by its high accuracy and optimal learning patterns before overfitting occurred. Overall, the sentiment analysis indicates that public opinion on TikTok regarding the topic under study tends to be positive at 72.90 percent, although there is still a portion of negative sentiment amounting to 27.10 percent.

5. References

- Abdi, A., Mariyam, S., Hasan, S., & Piran, J. (2019). Deep learning-based sentiment classification of evaluative text based on Multi-feature fusion. *Information Processing and Management*, 56(4), 1245–1259. <https://doi.org/10.1016/j.ipm.2019.02.018>
- Ahmad, W., Khan, H. U., & Alreshoodi, M. (2025). Aspect-Based Sentiment Analysis: A Comprehensive Review and Open Research Challenges. 13(February).
- Analysis, A. S. (2024). Examining Turkish Youth's Reaction to Instagram Influencers: July, 59–73.
- Ariyus, D., & Manongga, D. (2024). Enhancing Sentiment Analysis of Indonesian Tourism Video Content Commentary on TikTok: A FastText and Bi-LSTM Approach. 14(6), 18020–18028.
- Besman, A., Adiputra, A. V., & Saputra, S. J. (2018). Komunikasi Presiden Joko Widodo Dalam Vlog # Baliaman President Joko Widodo Communication In The # Baliaman. 21(1), 43–56. <https://doi.org/10.20422/jpk.v21i1.518>
- Cas, A. (2023). A review on sentiment analysis from social media platforms Margarita Rodríguez-Ib a. 223(March).
- Gray, J. E. (2021). The geopolitics of 'platforms': the TikTok challenge. 10(August 2020).
- Hussain, Z., Khan, A., & Ali, A. (2023). The Impact of User-Generated Content, Social Interactions and Virtual Economies on Metaverse Environments. 01(02).
- Kencana, W. H., Vicenovie, I., Situmeang, O., & Januar, K. (2024). TikTok as Digital Literacy Media for the 2024 Election in Indonesia. 1(1).
- Knittel, J., Koch, S., Tang, T., Chen, W., Wu, Y., Liu, S., & Ertl, T. (2021). Real-Time Visual Analysis of High-Volume Social Media Posts.
- Lestari, M. (2025). Sentiment Analysis of TikTok User Comments on Student Proposal Hearing Videos. 3(4), 147–152.
- Liu, B. (2012). Sentiment Analysis and Opinion Mining. May.
- Liyih, A., Anagaw, S., Yibeyin, M., & Tehone, Y. (2024). Sentiment analysis of the Hamas - Israel war on YouTube comments using deep learning. 1–9.
- Mutmatinah, S. (2020). Metode Deep Learning LSTM dalam Analisis Sentimen Aplikasi PeduliLindungi. <https://doi.org/10.34304/scientific.v1i1.231>
- Naratif, A., Publik, O., Isu, P., Thrifting, L., Media, D. I., & Twitter, S. (2023). Cyberspace sebagai ruang publik baru: analisis naratif opini publik pada isu larangan thrifting di media sosial twitter. 02(01), 1–25.
- Norhidayati, M. (2021). Impression Flow Based on Comment in Islamic Studies from Instagram Using Sentiment Analysis. 20(1), 42–56.
- Nurjanah, S., & Apidana, Y. H. (2025). Analisis Sentimen TikTok untuk Mengevaluasi Reputasi Merek Pasca Kasus Overclaim: Studi pada Daviena Skincare. 4(2).
- Pusat, B., Tanaman, S., Kong, H., Penelitian, B. B., & Pertanian, P. P. (2018). Deteksi Kecacatan Permukaan Buah Manggis Menggunakan Metode Deep Learning dengan Konvolusi Multilayer. 21(2), 230–236. <https://doi.org/10.18196/st.212229>
- Tika, L., Amelia, D., & Balqis, N. R. (2023). Changes in Communication Patterns in the Digital Age. 3(4), 544–556.
-

-
- Wei, Z. (2025). Cross-Platform Short-Video Diplomacy: Topic and Sentiment Analysis of China-US Relations on Douyin and TikTok.
- Wismashanti, R. A., & Indonesia, U. (2023). Social Media Content Moderation Challenges For Vulnerable Groups: A Case Study On Tiktok Indonesia. 3(8), 1449–1464.
- Wu, J., & Hong, T. (2022). The Picture of # Mentalhealth on Instagram: Congruent vs. Incongruent Emotions in Predicting the Sentiment of Comments. 7(April), 1–9.
<https://doi.org/10.3389/fcomm.2022.824119>
- Zhang, Y. (2023). Research on Sentiment Analysis of Government Short Video Comments based on BERT' s Multi Strategy Combination Model. 3(3).