Enhancing Sentiment Analysis of Indonesian Tourism Video Content Commentary on TikTok: A FastText and Bi-LSTM Approach

Dony Ariyus
Faculty of Computer Science, Amikom University of Yogyakarta, Indonesia
dony.a@amikom.ac.id (corresponding author)

Danny Manongga
Faculty of Information Technology, Satya Wacana Christian University, Salatiga, Indonesia
danny.manongga@uksw.edu

Irwan Sembiring
Faculty of Information Technology, Satya Wacana Christian University, Salatiga, Indonesia
irwan@uksw.edu

Received: 30 August 2024 | Revised: 19 September 2024 and 30 September 2024 | Accepted: 4 October 2024

 DOI: https://doi.org/10.48084/etasr.8859

ABSTRACT

Sentiment analysis is a method used to measure public opinion or the emotions of a group of people with similar interests based on their reactions to an event through text, images, videos, or audio on social media. However, such online data presents several challenges that can hinder the sentiment analysis process. These challenges stem mainly from the freedom that users have to post their content. Additionally, irrelevant opinions, often referred to as fake opinions, can also arise. The Bi-LSTM approach processes input sequences bidirectionally, allowing the model to capture information from both previous and subsequent contexts. This method is well-suited for sentiment analysis tasks due to its ability to recognize language nuances and relationships between different parts of the text. This study integrates a Bi-LSTM model with FastText word embeddings to filter out irrelevant opinions considered spam. The dataset consists of 150,351 TikTok comments taken from 100 popular videos related to tourist attractions. The experimental results show that the proposed Bi-LSTM model outperforms other models such as LSTM, CNN, GRU, MD-LSTM, and Peephole LSTM, achieving a test accuracy of 89.18%. Furthermore, when slang word translation is performed to convert slang into formal words, the Bi-LSTM model shows further improvement, with test accuracy reaching 93.10%, again surpassing the baseline models. These results demonstrate the robustness of the proposed method in handling noisy and informal language, thus improving the accuracy of sentiment analysis in the context of social media. This study provides a foundation for future research to improve sentiment analysis by addressing domain-specific challenges such as data imbalance and noise in social media data.

Keywords-social media data; sentiment analysis; Fasttext; Bi-LSTM

I. INTRODUCTION

Social media users have surged, particularly during the COVID-19 lockdown [1], as social media have transformed communication and the way information is spread around the world. Social media has redefined news creation and sharing, fostering dynamic interactions [2]. Platforms such as TikTok have gained popularity, especially among younger audiences, providing unique features [3]. TikTok serves as a powerful tool for gauging public sentiment on topics such as climate change, tourism, and food. However, the unstructured nature of user-generated content presents challenges in accurately capturing sentiment, especially in tourism. The rise of TikTok has also prompted organizations to adapt their strategies to engage younger audiences [4], while platforms like Douyin have become key in driving consumer activism and behavior [5].

Social media offer a space for experience-sharing, especially for young people, allowing them to express their attitudes and perceptions [6]. However, the diversity and informality of online opinions, often expressed in slang, complicate sentiment analysis. In tourism, social networks play a vital role in the promotion of destinations, allowing Destination Management Organizations (DMOs) to influence consumers through user-generated content [7]. However, its unstructured nature makes it difficult to analyze tourist feedback and draw actionable insights. TikTok is crucial in destination marketing, allowing users to share experiences that shape destination perception [8, 9]. TikTok food videos enhance tourists' focus on destination reputation, particularly around food quality and local context [10]. However, the unfiltered nature of comments complicates the extraction of useful insights for tourism organizations.

A key feature of TikTok is its commentary system, where users leave comments ranging from simple reactions to elaborated statements, often using emojis and informal language. This variety complicates sentiment analysis, as traditional methods struggle to capture the full context. By analyzing these comments, tourism organizations gain insight into audience perceptions and refine marketing strategies. Sentiment analysis helps identify areas for improvement [11, 12], improve tourist experiences, and support industry growth. Sentiment analysis is a key method for understanding social media content. However, managing informal language and imbalanced sentiment distributions poses challenges, especially on platforms such as TikTok. In tourism, sentiment analysis is critical for DMOs to monitor and interpret tourists' perceptions [13]. These insights help to develop targeted marketing strategies to enhance visitor experiences, requiring models capable of handling noisy, unstructured data while delivering accurate sentiment classification [14].

Sentiment analysis uses Natural Language Processing (NLP) to categorize social media content as positive, negative, or neutral by examining the language used [15, 16]. Common approaches include lexicon-based techniques and machine-learning methods [17]. Lexicon-based methods [18] rely on predefined word lists with sentiment ratings. These methods are efficient, but often fail to capture language nuances, especially in social media, where slang and informal language are common. To improve accuracy, lexicon-based methods are frequently combined with machine learning techniques [19]. Machine learning methods such as FastText and Bi-LSTM have gained prominence for handling the complexities of social media language. FastText captures the semantic meaning of words, even slang, while Bi-LSTM processes sequential data and captures long-range contextual relationships [20]. Bi-LSTM analyzes input in both directions, making it particularly well-suited for sentiment analysis [21]. By combining FastText and Bi-LSTM, the proposed method addresses challenges such as informal language, imbalanced sentiment categories, and slang, making it a promising approach for analyzing tourism-related social media content [23].

In [23], sentiments were analyzed in visual content at renowned destinations, such as St. Mark's Square and Doge's Palace, involving Tripadvisor comments from 2023 and logistic regression to identify factors influencing reviews. However, this study has limitations, as sentiment analysis cannot yet identify weather conditions or nearby attractions. In [21], Twitter sentiment data were analyzed. However, limitations remain, as current methods cannot yet detect weather conditions or nearby attractions. A study on Twitter sentiment related to tourism in Thailand during COVID-19 compared the CART, Random Forest, and SVM methods, with the latter being the most accurate. Oversampling produced better results than undersampling, and unigrams outperformed bigrams [24]. These findings suggest that machine learning algorithms can effectively predict sentiment and tourist intentions, with word usage in each class playing a key role in identifying sentiments and intentions to visit Thailand. In [25], research on 39,216 TripAdvisor reviews and additional datasets analyzed the negative reputation of Marrakech tourist attractions. Latent Dirichlet Allocation (LDA) and a lexicon-based approach were used to extract hidden aspects of tourist feedback and identify weaknesses, enhancing the tourist experience.

LSTM has become essential for analyzing sentiment in industries such as healthcare and tourism. In [22], an LSTM-based social media sentiment analysis achieved 81.15% accuracy, outperforming other machine learning algorithms. Other machine learning techniques, such as logistic regression, naive Bayes, XGBoost, and random forest, have been applied to rainfall prediction. In a dataset of 145,460 records, XGBoost achieved the highest accuracy of 84.61% [26]. A study comparing CNN, LSTM, CNN-LSTM, BiLSTM, and ConvBiLSTM using TripAdvisor hotel reviews found that LSTM was the most accurate (96.42%). In [27], CNN, LSTM, CNN-LSTM, BiLSTM, and ConvBiLSTM were compared on TripAdvisor hotel reviews, with LSTM achieving the highest accuracy (96.42%). In [28], AraWord2Vec was developed, which is a custom word embedding model, and CM_BiLSTM achieved 98.47% accuracy in binary classification and 98.92% in multiclass classification for Arabic sentiment analysis.

This study proposes a method for sentiment analysis tailored to Indonesian super-priority tourism destinations using FastText and Bi-LSTM. A substantial dataset of TikTok videos was collected, which were related to these destinations, and the data were preprocessed by tokenizing, stemming, and removing stop words. Then, the proposed sentiment analysis model was trained and benchmarked against other state-of-the-art techniques. The research objectives of this study were:

- Combine FastText and Bi-LSTM for sentiment analysis, to overcome the challenges of informal language, slang, and irregular data structures in social media, and achieve better accuracy than traditional methods.
- Apply the proposed model to analyze the TikTok comments related to super-priority tourist destinations in Indonesia, providing useful insights for DMOs and policymakers.
- Compare the performance of the proposed with other methods, such as CNN, LSTM, MD-LSTM, Peephole LSTM, and GRU. The results show superior accuracy and ability to handle noisy data for the proposed method.

II. PROPOSED METHOD

The proposed method consists of five stages: preprocessing, labeling, word embedding, modeling, and evaluation. First, commentary from TikTok videos featuring Indonesia's super-priority tourist destinations was collected. These videos were officially uploaded by the Indonesian Ministry of Tourism. The method interpreted and analyzed TikTok comments, including those mentioning other users. The method process is illustrated in Figure 1.

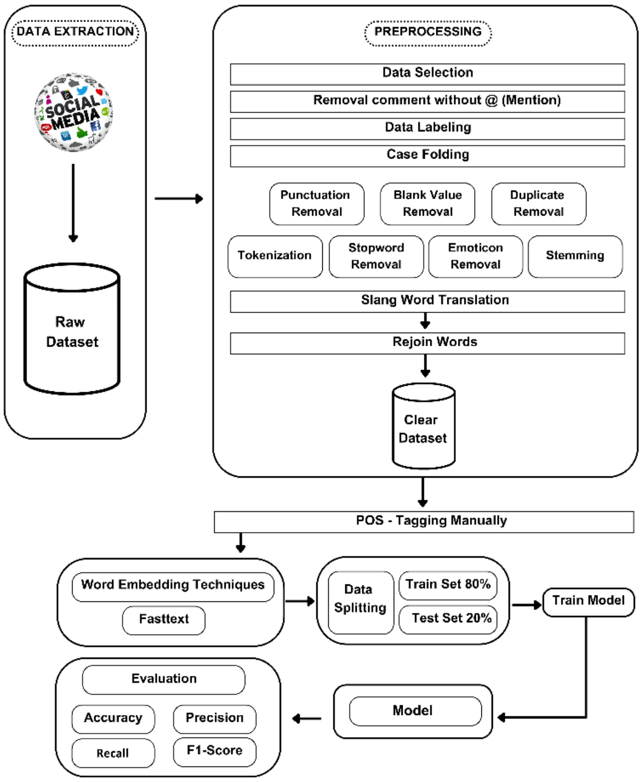

Fig. 1. The proposed method.

Each stage consists of multiple substages. For example, in the preprocessing stage, the raw data undergo case-folding, comment filtering, and rejoining words.

A. *Preprocessing*

Data preprocessing [29, 30] in sentiment analysis of tourist destinations involves the processes of cleaning, transforming, and preparing textual data related to tourist reviews or comments about specific destinations before applying sentiment analysis techniques. Preprocessing aims to improve the quality and usability of the data by eliminating noise, standardizing the formats, and extracting pertinent information. Preprocessing consisted of the following substages [31]: case folding, commentary filtering, punctuation removal, duplicate removal, empty, stopword removal, emoticon removal, tokenization, stemming, slang word translation, and rejoining. Preprocessing is vital in sentiment analysis as it prepares the text data for subsequent analysis and model development. The different preprocessing steps applied to the input text improve data quality and boost the effectiveness of sentiment analysis models. Once the data preprocessing is done, the data can be labeled for further processing.

B. *Data Labeling*

Data labeling, or annotation or tagging, refers to assigning meaningful and informative labels or tags to data instances. In sentiment analysis, data labeling involves manually or automatically assigning sentiment labels to textual data [32]. Data labeling is a critical step in supervised machine learning, as it determines the sentiment conveyed in the data, such as positive, negative, or neutral, and the labeled data are used to train models to predict sentiment or perform other tasks accurately. Annotated data serve as the ground truth or reference against which the model learns to make predictions. This study used two approaches to data labeling: manual labeling and post-tagging.

Manual labeling [31, 32] is when human annotators carefully analyze text data and assign sentiment labels based on predefined guidelines. Through language proficiency, contextual understanding, and subjective judgment, the annotators determine the sentiment expressed in each instance, considering linguistic cues and the overall tone. Manual labeling captures subtle sentiment patterns that automated methods may overlook, resulting in more precise and contextually nuanced sentiment analysis outcomes. Although it requires human resources and time, manual labeling ensures higher quality and reliability by incorporating human understanding and interpretation of sentiment.

Post-tagging involves automatically assigning sentiment labels to textual data using pre-existing or pre-trained models or algorithms [33]. Instead of relying on manual annotation, post-tagging utilizes machine learning methods to analyze text and predict the conveyed sentiment. Through the learned patterns and features, the model assigns sentiment labels, such as positive, negative, or neutral, to data samples. Post-tagging enhances efficiency and scalability by eliminating the requirement for human annotators. However, the accuracy and reliability of post-tagging are critically dependent on the performance and quality of the sentiment analysis model or algorithm employed.

C. *Word Embedding*

Word embedding [34] is an NLP technique that represents words or phrases as dense and continuous functions mapped to a high-dimensional vector space. Its purpose is to analyze the way words relate to each other and how their meanings are influenced by the surrounding text based on their usage in a given text corpus. Word embedding models are trained on vast amounts of text using neural networks or machine learning [35-39]. In a word embedding model, each word is represented by a fixed-length vector, typically consisting of several hundred dimensions. A word's vector representation is learned by analyzing how it co-occurs with other words in the training data. Words that appear in similar contexts or share similar meanings usually have comparable vector representations, placing them closer together in the embedding space. The benefit of word embeddings is that they capture both semantic and syntactic relationships between words. For example, words related in meaning, such as "king" and "queen," are expected to have similar vector representations and therefore be close in the embedding space. This allows algorithms to capture the meaning of words and their associations, enabling more effective analysis and processing of text data. The proposed method utilizes FastText for word embedding.

FastText [40-42] is a word embedding approach developed by Facebook AI Research. It is an extension of the well-known word2vec algorithm that not only captures word-level representations but also incorporates subword information.

FastText represents words as vectors by learning to predict the likelihood of a word given a certain context, its context, or vice versa. The FastText approach decomposes words into sub-word units known as character n-grams. For example, the word "apple" can be represented by its character n-grams: "ap," "app," "ppl," "ple," and "le." By considering these subword units, FastText can capture morphological and syntactic information, making it particularly useful for languages with rich morphology or when dealing with out-of-vocabulary words. FastText constructs a dictionary of words and their sub-word units during the training process. It then learns to assign vector representations to both words and subwords based on their co-occurrence patterns within a given text corpus [43]. The resulting word vectors can capture semantic relationships between words and their sub-word components.

D. Modeling

Modeling in sentiment analysis involves developing and training machine or deep learning models to predict the sentiment or emotional polarity of textual data [44]. The goal is to create a model that accurately classifies text into positive, negative, or neutral sentiment categories based on its content and context. The proposed method implements Bi-LSTM for modeling [45, 46]. Bi-LSTM stands for Bidirectional Long Short-Term Memory, a type of Recurrent Neural Network (RNN) architecture frequently used in NLP tasks such as sentiment analysis. Bi-LSTM integrates both forward and backward information flows, enabling it to capture contextual information from both previous and subsequent words in a sequence. In a traditional LSTM, information flows sequentially from the input to the output in one direction, typically from left to right. Bi-LSTM introduces an additional layer that processes the input sequence in the reverse direction, from right to left. This bidirectional flow allows the model to capture dependencies and contextual information from both preceding and following words in the sequence, resulting in a more comprehensive understanding of the input text. The output of the Bi-LSTM can subsequently be used for additional tasks, such as classification or sentiment prediction. It is important to note that the computations in an LSTM, including Bi-LSTM, involve various mathematical operations such as matrix multiplications, element-wise operations, and activation functions (e.g., sigmoid and tanh). These computations are performed to update and propagate information through the LSTM layers, allowing the network to capture sequential dependencies and contextual information in the input text data. Figure 2 illustrates the structure of the Bi-LSTM network.

E. Evaluation

Evaluation involves measuring the accuracy, reliability, and quality of the predictions made by the model in classifying the sentiment expressed in the textual data. Several commonly used evaluation metrics exist in sentiment analysis:

- The confusion matrix is a table that displays four different combinations of predicted and actual values, representing the results of the classification process, including True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN). It helps identify specific errors the model might be making and offers insights into its performance across various sentiment categories.
- Recall, also known as sensitivity or the true positive rate, describes the model's success in accurately identifying and retrieving relevant information. It reflects the model's ability to identify all pertinent sentiment events.

$$Recall = \frac{TP}{TP+FN} \quad (1)$$

- Precision defines the rate of correct positives to all positive predictions.

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

- Accuracy measures how accurately the model predicts the sentiment made by the model in the available dataset.

$$Accuracy = \frac{TP+FN}{TP+TN+FP+FN} \quad (3)$$

- The F1 score evaluates the weighted average of precision and recall, offering a balanced measure of model performance. Symmetric results indicate a better balance between precision and recall.

$$F1-score = 2 \times \frac{Precision \times Recall}{Precision+Recall} \quad (4)$$

It is important to note that sentiment analysis evaluation relies heavily on having a labeled evaluation dataset with ground-truth sentiment labels. The dataset should represent the target domain and cover a range of sentiment expressions to assess the model's performance. By evaluating sentiment analysis models using appropriate metrics and techniques, researchers and practitioners can evaluate their models' strengths and weaknesses by identifying areas for improvement, enabling them to make informed choices in model selection and optimization.

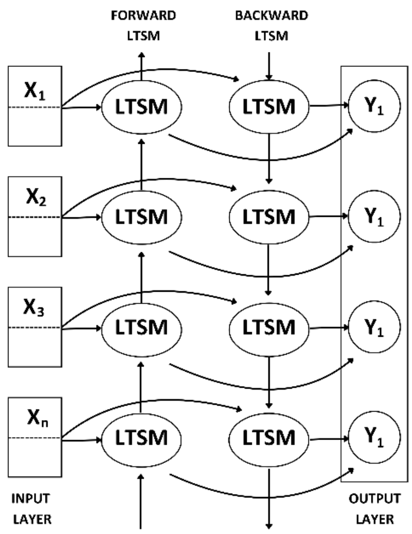

Fig. 2. The Bi-LSTM process captures sequential features, with the final hidden layer of the LSTM serving as the feature representation of the text.

III. RESULTS AND DISCUSSION

The dataset was sourced from the TikTok comment section of super-priority tourist destinations identified by the Ministry of Tourism of Indonesia. In the preprocessing stage, comments were selected if they mentioned another TikTok user. Then, the comment was extracted, and the mentioned username was ignored. This type of comment was chosen because it could influence the user mentioned on whether to visit the tourist destination. Additionally, this selection process reduced the number of comments available for analysis. A total of 150,351 comments were collected from 120 videos. Table I presents an example of the dataset containing comments scraped from TikTok videos that meet the data collection criteria.

TABLE I. TIKTOK COMMENT DATA

No	Comment Text
1	woowwww jadi pengen visit lagi ke LBajo, wkt itu gak sampai sini...
2	@tehgelaspakey @ariesssss07 yang sukses ya temen temen biar bisa ke bajo
3	Rasanya anda menjadi ironmen kalo kesana @giapuphiy @witadp_ @fayyzaazhra30
4	sini sih fix @ka.yiin @dhafinsydn @dkiak
5	@syatsyaa kuy

This study focuses on comments related to TikTok videos, identified using keywords associated with tourist destinations. The first step in data preprocessing was to filter comments containing the @ symbol and indicating interactions between users. These filtered comments were then used as the dataset for further processing. After the filtering process, the number of comments decreased significantly, as shown in Table II.

TABLE II. INITIAL COMPOSITION OF THE DATASET

Destination	Number of comments of Tiktok Video				
	Total comments	With @	Positive	Negative	Neutral
Labuah Bajo	150351	1687	936	123	628
Danau Toba		1248	664	145	439
Candi Borobudur		1663	875	290	498
Mandalika		1549	762	138	649

TABLE III. DATASET COMPOSITION AFTER NORMALIZATION

Destination	Number of comments				
	Total comments	With @	Positive	Negative	Neutral
Labuah Bajo	150351	2808	936	936	936
Danau Toba		1992	664	664	664
Candi Borobudur		2625	875	875	875
Mandalika		1947	649	649	649

The data consist of 936 positive, 123 negative, and 628 neutral comments. These data were labeled manually using the post-tagging technique. The dataset exhibits an imbalance for each tourist attraction video, which affects the model's ability to accurately identify the minority class during classification tasks. Random oversampling was used to address this by randomly replicating instances from the minority class, based on the dataset size and the specified oversampling rate.

However, various studies [27, 47, 48] suggest that random oversampling may increase the risk of overfitting, as it replicates the samples from the minority class exactly. Consequently, the proportion of minority labels (Neutral and Negative) is adjusted to match the number of the majority class (Positive). The proposed method embeds the data using FastText, which was chosen because it is considered more robust than the existing technique. The data was then randomly divided into three sets: training, testing, and validation. The train set represents 80% of the filtered dataset, the test set represents 20%, and the validation set represents 20% of the filtered dataset.

The proposed BI-LSTM and Fasttext technique was compared with six state-of-the-art methods: modified LSTM from SEE, CNN [49], GRU [50], MD-LSTM [51], Peephole LSTM [52], and Stacked LSTM [42]. The experimental results, as presented in Tables III and IV, clearly demonstrate that the proposed method outperformed the others.

TABLE IV. ACCURACY OF COMPARED METHODS

Method	Training accuracy	Validation accuracy	Testing accuracy
LSTM	0.9450	0.8898	0.8795
Bi-LSTM (Proposed)	**0.9489**	**0.8891**	**0.8918**
Convolutional Neural Networks (CNN)	0.9471	0.8879	0.8849
Gated Recurrent Unit (GRU)	0.9443	0.8921	0.8723
Multidimensional LSTM (MD-LSTM)	0.9400	0.8804	0.8804
Peephole LSTM	0.9449	0.8894	0.8795
Stacked LSTM	0.9350	0.8827	0.8711

TABLE V. ACCURACY BETWEEN THE COMPARED METHODS WITH SLANGWORD TRANSLATION

Method	Training accuracy	Validation accuracy	Testing accuracy
LSTM	0.9561	0.9398	0.9297
Bi-LSTM (Proposed)	**0.9612**	**0.9432**	**0.9310**
Convolutional Neural Networks (CNN)	0.9557	0.9368	0.9209
Gated Recurrent Unit (GRU)	0.9555	0.9374	0.9229
Multidimensional LSTM	0.9561	0.9361	0.9234
Peephole LSTM	0.9599	0.9389	0.9297
Stacked LSTM	0.9587	0.9380	0.9275

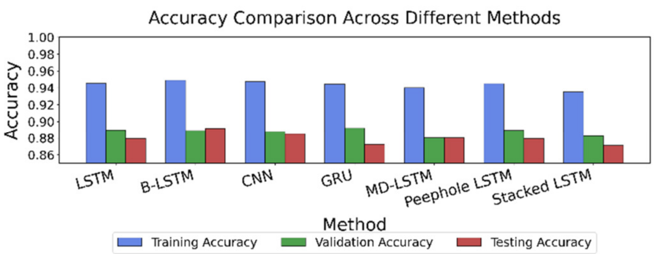

Fig. 3. Accuracy comparison between the proposed and existing methods.

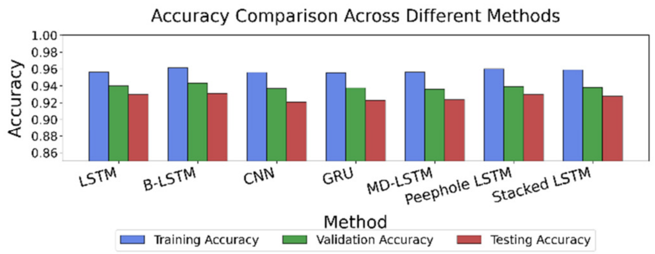

Fig. 4. Accuracy comparison between the proposed and existing methods with slangword translation.

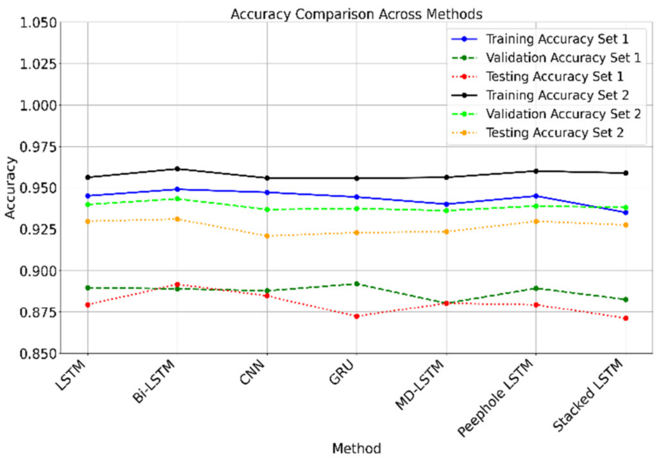

Fig. 5. Testing accuracy results of the six modeling algorithms.

By incorporating 20,000 words of Indonesian social media slang into the language model, there is a notable improvement in the performance of each algorithm. This infusion of colloquial expressions enhances the model's ability to understand and interpret informal language, ultimately refining its proficiency in various computational tasks. The inclusion of social media slang in the dataset contributes to a notable enhancement of 5% compared to the initial results. This significant performance improvement enhances the effectiveness of the Bi-LSTM model in processing and managing datasets in the Indonesian language.

IV. LIMITATIONS AND STRENGTHS

This study has several limitations that should be considered. First, this study aimed to support the super-priority tourist destinations designated by the Indonesian Ministry of Tourism, but the necessary dataset was not yet available. TikTok was chosen as the social media platform for data collection due to the uniqueness of its comments, which tend to be shorter and often use informal language. The collection of raw datasets and data preprocessing took almost one year. Testing various word embedding and machine learning methods did not yield optimal results, so three annotators were employed to ensure that sentiments were labeled accurately. Due to the prevalence of short and informal comments, a slang translation feature was added during data preprocessing to improve accuracy. The author compiled more than 20,000 Indonesian slang words to enhance the model's performance. For even better results in processing short texts and informal language, a larger collection of informal words is necessary.

V. CONCLUSION

This study endeavored to contribute to the existing literature by examining comments on social media, specifically TikTok in Indonesian, about the super-priority tourism initiatives outlined by the Ministry of Tourism. Based on this objective, the study sought to explore public sentiment toward the most favored tourist destinations within the archipelago. Furthermore, the study aimed to identify a suitable machine-learning algorithm model for this context. Six different methods were applied and compared using a dataset comprising 150,351 TikTok comments. After preprocessing, this dataset was refined to 6,147 data points, adhering to the criterion that comments must include a mention (@) to offer insight into the sentiments (positive, negative, or neutral) expressed toward tourist attractions. The test results from the implementation of six methods fell below the 90% threshold, primarily due to a substantial portion of the data remaining unrecognized, possibly due to the usage of language not present in the corpus. To enhance test results, approximately 20,000 terms and slang were collected after preprocessing. Interestingly, all methods showed a 5% improvement with this enhancement. In particular, the Bi-LSTM method continued to demonstrate superiority over the other applied methods in this context.

In the Indonesian context, TikTok is extensively utilized by the youth, who exhibit a heightened inclination toward travel, particularly in their pursuit of visually appealing tourist destinations for photo opportunities. Expressing their recommendations through TikTok comments, these individuals predominantly employ a language that resonates well with the youth demographic. It is crucial to acknowledge the perpetual growth of slang and abbreviations on social media platforms. To address this, this study consistently gathered and incorporated these linguistic nuances into a corpus, thereby creating a resource for future researchers. This corpus is designed to enhance deep learning-based sentiment analysis on social media, ensuring a more nuanced understanding of user sentiments in this dynamic digital landscape.

The combination of FastText with Bi-LSTM has been proven to deliver better results compared to other methods in this dataset. FastText's ability to handle out-of-vocabulary words, along with Bi-LSTM's contextual learning capabilities, significantly enhances the model's accuracy.

VI. FUTURE WORK

Preprocessing is a crucial step in achieving more accurate results in sentiment analysis. Defining slang words can enhance accuracy by an average of 5% when combining machine learning methods with FastText in the tourism domain. The best method in this study can also be extended and applied to other domains and multilingual datasets to assess its effectiveness.

(a)

(b)

(c)

(d)

(e)

(f)

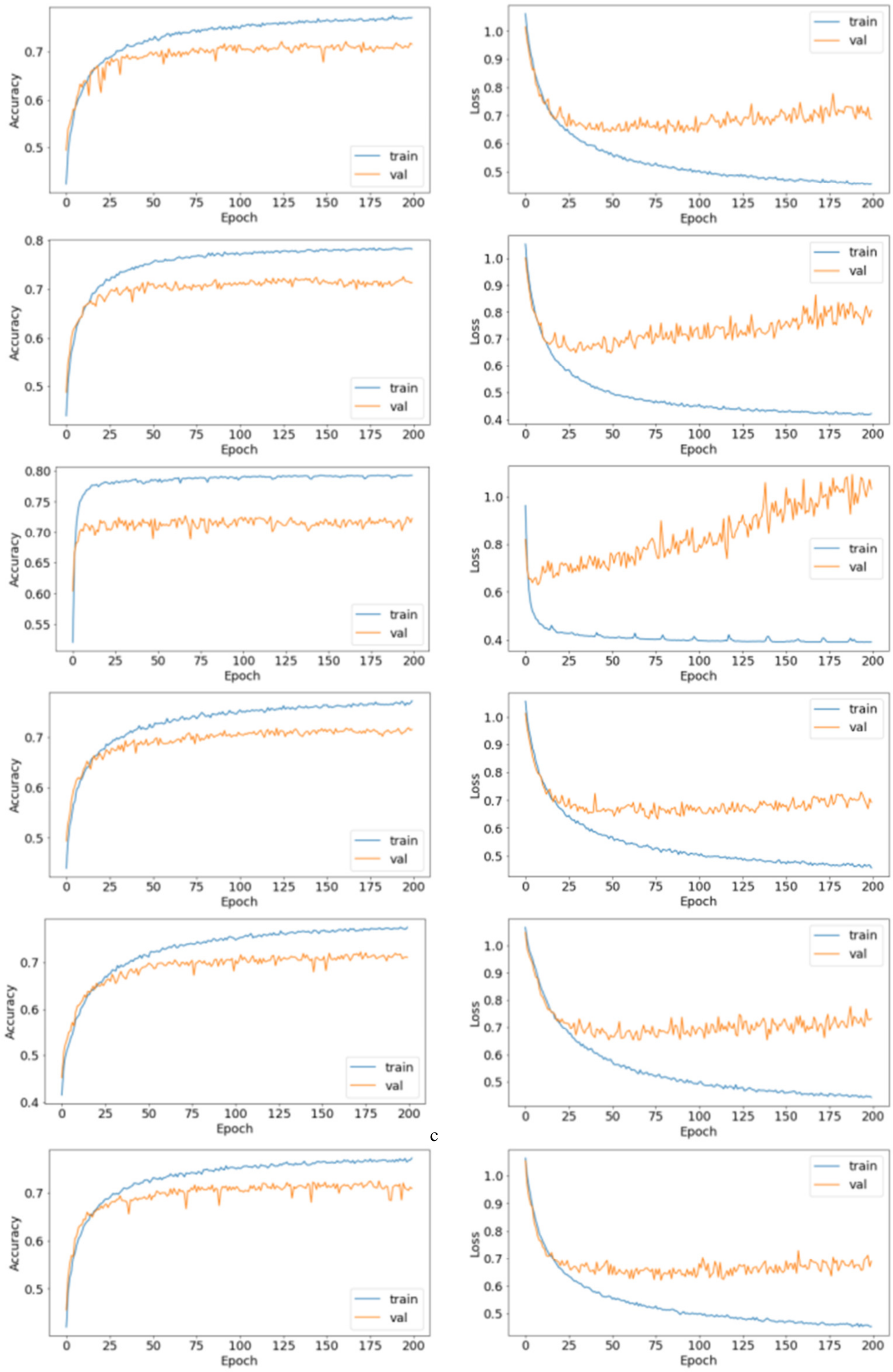

Fig. 6. Training accuracy and loss of the six methods compared: (a) LTSM, (b) Bi-LSTM, (c) CNN, (d), GRU, (e) MD-LSTM, (f) Peephole LSTM.

REFERENCES

[1] C. H. Basch, B. Yalamanchili, and J. Fera, "#Climate Change on TikTok: A Content Analysis of Videos," *Journal of Community Health*, vol. 47, no. 1, pp. 163–167, Feb. 2022, https://doi.org/10.1007/s10900-021-01031-x.

[2] M. C. Negreira-Rey, J. Vázquez-Herrero, and X. López-García, "Blurring Boundaries Between Journalists and Tiktokers: Journalistic Role Performance on TikTok," *Media and Communication*, vol. 10, no. 1, pp. 146–156, Feb. 2022, https://doi.org/10.17645/mac.v10i1.4699.

[3] P. Cuesta-Valiño, P. Gutiérrez-Rodríguez, and P. Durán-Álamo, "Why Do People Return to Video Platforms? Millennials and Centennials on TikTok," *Media and Communication*, vol. 10, no. 1, pp. 198–207, Feb. 2022, https://doi.org/10.17645/mac.v10i1.4737.

[4] J. Vázquez-Herrero, M. C. Negreira-Rey, and A. I. Rodríguez-Vázquez, "Intersections between TikTok and TV: Channels and Programmes Thinking Outside the Box," *Journalism and Media*, vol. 2, no. 1, pp. 1–13, Mar. 2021, https://doi.org/10.3390/journalmedia2010001.

[5] Z. Yu, J. Hou, and O. T. Zhou, "Short Video Activism With and on Douyin: An Innovative Repertoire of Contention for Chinese Consumers," *Social Media + Society*, vol. 9, no. 1, Jan. 2023, Art. no. 20563051231157603, https://doi.org/10.1177/20563051231157603.

[6] C. Are and P. Briggs, "The Emotional and Financial Impact of De-Platforming on Creators at the Margins," *Social Media + Society*, vol. 9, no. 1, Jan. 2023, Art. no. 20563051231155103, https://doi.org/10.1177/20563051231155103.

[7] I. Hipiny, H. Ujir, A. A. Alias, M. Shanat, and M. K. Ishak, "Who danced better? ranked tiktok dance video dataset and pairwise action quality assessment method," *International Journal of Advances in Intelligent Informatics*, vol. 9, no. 1, pp. 96–107, Mar. 2023, https://doi.org/10.26555/ijain.v9i1.919.

[8] Y. Li, X. Xu, B. Song, and H. He, "Impact of Short Food Videos on the Tourist Destination Image—Take Chengdu as an Example," *Sustainability*, vol. 12, no. 17, Jan. 2020, Art. no. 6739, https://doi.org/10.3390/su12176739.

[9] C. Zhu, L. H. N. Fong, H. Gao, and C. Y. N. Liu, "When TikTok meets celebrity: an investigation of how celebrity attachment influences visit intention," *Current Issues in Tourism*, vol. 26, no. 17, pp. 2762–2776, Sep. 2023, https://doi.org/10.1080/13683500.2022.2097058.

[10] X. Wang, Y. Yu, Z. Zhu, and J. Zheng, "Visiting Intentions toward Theme Parks: Do Short Video Content and Tourists' Perceived Playfulness on TikTok Matter?," *Sustainability*, vol. 14, no. 19, Jan. 2022, Art. no. 12206, https://doi.org/10.3390/su141912206.

[11] M. R. R. Rana, A. Nawaz, T. Ali, A. M. El-Sherbeeny, and W. Ali, "A BiLSTM-CF and BiGRU-based Deep Sentiment Analysis Model to Explore Customer Reviews for Effective Recommendations," *Engineering, Technology & Applied Science Research*, vol. 13, no. 5, pp. 11739–11746, Oct. 2023, https://doi.org/10.48084/etasr.6278.

[12] N. Sureja, N. Chaudhari, P. Patel, J. Bhatt, T. Desai, and V. Parikh, "Hyper-tuned Swarm Intelligence Machine Learning-based Sentiment Analysis of Social Media," *Engineering, Technology & Applied Science Research*, vol. 14, no. 4, pp. 15415–15421, Aug. 2024, https://doi.org/10.48084/etasr.7818.

[13] A. Hussain and E. Cambria, "Semi-supervised learning for big social data analysis," *Neurocomputing*, vol. 275, pp. 1662–1673, Jan. 2018, https://doi.org/10.1016/j.neucom.2017.10.010.

[14] Z. Abbasi-Moud, H. Vahdat-Nejad, and J. Sadri, "Tourism recommendation system based on semantic clustering and sentiment analysis," *Expert Systems with Applications*, vol. 167, Apr. 2021, Art. no. 114324, https://doi.org/10.1016/j.eswa.2020.114324.

[15] N. A. Alabdulkarim, M. A. Haq, and J. Gyani, "Exploring Sentiment Analysis on Social Media Texts," *Engineering, Technology & Applied Science Research*, vol. 14, no. 3, pp. 14442–14450, Jun. 2024, https://doi.org/10.48084/etasr.7238.

[16] A. S. Imran, S. M. Daudpota, Z. Kastrati, and R. Batra, "Cross-Cultural Polarity and Emotion Detection Using Sentiment Analysis and Deep Learning on COVID-19 Related Tweets," *IEEE Access*, vol. 8, pp. 181074–181090, 2020, https://doi.org/10.1109/ACCESS.2020.3027350.

[17] A. Alsayat, "Improving Sentiment Analysis for Social Media Applications Using an Ensemble Deep Learning Language Model," *Arabian Journal for Science and Engineering*, vol. 47, no. 2, pp. 2499–2511, Feb. 2022, https://doi.org/10.1007/s13369-021-06227-w.

[18] A. Ishaq, S. Asghar, and S. A. Gillani, "Aspect-Based Sentiment Analysis Using a Hybridized Approach Based on CNN and GA," *IEEE Access*, vol. 8, pp. 135499–135512, 2020, https://doi.org/10.1109/ACCESS.2020.3011802.

[19] R.Chundi, V. R. Hulipalled, and J. B. Simha, "NBLex: emotion prediction in Kannada-English code-switch text using naïve bayes lexicon approach," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 13, no. 2, pp. 2068–2077, Apr. 2023.

[20] R. A. Stein, P. A. Jaques, and J. F. Valiati, "An analysis of hierarchical text classification using word embeddings," *Information Sciences*, vol. 471, pp. 216–232, Jan. 2019, https://doi.org/10.1016/j.ins.2018.09.001.

[21] M. Fattah and M. A. Haq, "Tweet Prediction for Social Media using Machine Learning," *Engineering, Technology & Applied Science Research*, vol. 14, no. 3, pp. 14698–14703, Jun. 2024, https://doi.org/10.48084/etasr.7524.

[22] H. Jelodar, Y. Wang, R. Orji, and S. Huang, "Deep Sentiment Classification and Topic Discovery on Novel Coronavirus or COVID-19 Online Discussions: NLP Using LSTM Recurrent Neural Network Approach," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 10, pp. 2733–2742, Jul. 2020, https://doi.org/10.1109/JBHI.2020.3001216.

[23] E. Bigne, C. Ruiz, A. Cuenca, C. Perez, and A. Garcia, "What drives the helpfulness of online reviews? A deep learning study of sentiment analysis, pictorial content and reviewer expertise for mature destinations," *Journal of Destination Marketing & Management*, vol. 20, Jun. 2021, Art. no. 100570, https://doi.org/10.1016/j.jdmm.2021.100570.

[24] N. Leelawat *et al.*, "Twitter data sentiment analysis of tourism in Thailand during the COVID-19 pandemic using machine learning," *Heliyon*, vol. 8, no. 10, Oct. 2022, https://doi.org/10.1016/j.heliyon.2022.e10894.

[25] T. Ali, B. Marc, B. Omar, K. Soulaimane, and S. Larbi, "Exploring destination's negative e-reputation using aspect based sentiment analysis approach: Case of Marrakech destination on TripAdvisor," *Tourism Management Perspectives*, vol. 40, Oct. 2021, Art. no. 100892, https://doi.org/10.1016/j.tmp.2021.100892.

[26] R. A. Hasan, M. F. Alomari, and J. B. Jamaluddin, "Comparative study: Using machine learning techniques about rainfall prediction," *AIP Conference Proceedings*, vol. 2787, no. 1, Jul. 2023, Art. no. 050014, https://doi.org/10.1063/5.0148472.

[27] F. Amali, H. Yigit, and Z. H. Kilimci, "Sentiment Analysis of Hotel Reviews using Deep Learning Approaches," in *2024 IEEE Open Conference of Electrical, Electronic and Information Sciences (eStream)*, Vilnius, Lithuania, Apr. 2024, pp. 1–8, https://doi.org/10.1109/eStream61684.2024.10542593.

[28] M. Abbes, Z. Kechaou, and A. M. Alimi, "A Novel Hybrid Model Based on CNN and Bi-LSTM for Arabic Multi-domain Sentiment Analysis," in *Complex, Intelligent and Software Intensive Systems*, 2023, pp. 92–102, https://doi.org/10.1007/978-3-031-35734-3_10.

[29] M. Mujahid *et al.*, "Sentiment Analysis and Topic Modeling on Tweets about Online Education during COVID-19," *Applied Sciences*, vol. 11, no. 18, 2021, https://doi.org/10.3390/app11188438.

[30] A. Alsaeedi and M. Zubair, "A Study on Sentiment Analysis Techniques of Twitter Data," *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 2, 2019, https://doi.org/10.14569/IJACSA.2019.0100248.

[31] F. Iqbal *et al.*, "A Hybrid Framework for Sentiment Analysis Using Genetic Algorithm Based Feature Reduction," *IEEE Access*, vol. 7, pp. 14637–14652, 2019, https://doi.org/10.1109/ACCESS.2019.2892852.

[32] N. A. K. M. Haris, S. Mutalib, A. M. A. Malik, S. Abdul-Rahman, and S. N. K. Kamarudin, "Sentiment classification from reviews for tourism analytics," *International Journal of Advances in Intelligent Informatics*, vol. 9, no. 1, pp. 108–120, Mar. 2023, https://doi.org/10.26555/ijain.v9i1.1077.

[33] M. Araújo, A. Pereira, and F. Benevenuto, "A comparative study of machine translation for multilingual sentence-level sentiment analysis," *Information Sciences*, vol. 512, pp. 1078–1102, Feb. 2020, https://doi.org/10.1016/j.ins.2019.10.031.

[34] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, "Distributed Representations of Words and Phrases and their Compositionality," in *Advances in Neural Information Processing Systems*, 2013, vol. 26.

[35] S. Rida-E-Fatima *et al.*, "A Multi-Layer Dual Attention Deep Learning Model With Refined Word Embeddings for Aspect-Based Sentiment

Analysis," *IEEE Access*, vol. 7, pp. 114795–114807, 2019, https://doi.org/10.1109/ACCESS.2019.2927281.

[36] M. Ghorbani, M. Bahaghighat, Q. Xin, and F. Özen, "ConvLSTMConv network: a deep learning approach for sentiment analysis in cloud computing," *Journal of Cloud Computing*, vol. 9, no. 1, Mar. 2020, Art. no. 16, https://doi.org/10.1186/s13677-020-00162-1.

[37] P. Sánchez-Núñez, M. J. Cobo, C. D. L. Heras-Pedrosa, J. I. Peláez, and E. Herrera-Viedma, "Opinion Mining, Sentiment Analysis and Emotion Understanding in Advertising: A Bibliometric Analysis," *IEEE Access*, vol. 8, pp. 134563–134576, 2020, https://doi.org/10.1109/ACCESS.2020.3009482.

[38] S. A. M. Vermeer, T. Araujo, S. F. Bernritter, and G. van Noort, "Seeing the wood for the trees: How machine learning can help firms in identifying relevant electronic word-of-mouth in social media," *International Journal of Research in Marketing*, vol. 36, no. 3, pp. 492–508, Sep. 2019, https://doi.org/10.1016/j.ijresmar.2019.01.010.

[39] I. Portugal, P. Alencar, and D. Cowan, "The use of machine learning algorithms in recommender systems: A systematic review," *Expert Systems with Applications*, vol. 97, pp. 205–227, May 2018, https://doi.org/10.1016/j.eswa.2017.12.020.

[40] K. Fiok, W. Karwowski, E. Gutierrez, and M. Reza-Davahli, "Comparing the Quality and Speed of Sentence Classification with Modern Language Models," *Applied Sciences*, vol. 10, no. 10, May 2020, Art. no. 3386, https://doi.org/10.3390/app10103386.

[41] Y. Xu, S. Chen, and X. Xu, "Research on Viewpoint Extraction in Microblog," *Intelligent Automation & Soft Computing*, vol. 29, no. 3, pp. 495–511, 2021, https://doi.org/10.32604/iasc.2021.018896.

[42] A. Onan and M. A. Toçoğlu, "A Term Weighted Neural Language Model and Stacked Bidirectional LSTM Based Framework for Sarcasm Identification," *IEEE Access*, vol. 9, pp. 7701–7722, 2021, https://doi.org/10.1109/ACCESS.2021.3049734.

[43] A. Ezen-Can, "A Comparison of LSTM and BERT for Small Corpus." arXiv, Sep. 11, 2020, https://doi.org/10.48550/arXiv.2009.05451.

[44] E. Saquete, J. Zubcoff, Y. Gutiérrez, P. Martínez-Barco, and J. Fernández, "Why are some social-media contents more popular than others? Opinion and association rules mining applied to virality patterns discovery," *Expert Systems with Applications*, vol. 197, Jul. 2022, Art. no. 116676, https://doi.org/10.1016/j.eswa.2022.116676.

[45] U. Naseem, I. Razzak, M. Khushi, P. W. Eklund, and J. Kim, "COVIDSenti: A Large-Scale Benchmark Twitter Data Set for COVID-19 Sentiment Analysis," *IEEE Transactions on Computational Social Systems*, vol. 8, no. 4, pp. 1003–1015, Dec. 2021, https://doi.org/10.1109/TCSS.2021.3051189.

[46] N. J. Prottasha *et al.*, "Transfer Learning for Sentiment Analysis Using BERT Based Supervised Fine-Tuning," *Sensors*, vol. 22, no. 11, 2022, https://doi.org/10.3390/s22114157.

[47] R. Tajvidi and A. Karami, "The effect of social media on firm performance," *Computers in Human Behavior*, vol. 115, Feb. 2021, Art. no. 105174, https://doi.org/10.1016/j.chb.2017.09.026.

[48] A. Setyanto *et al.*, "Arabic Language Opinion Mining Based on Long Short-Term Memory (LSTM)," *Applied Sciences*, vol. 12, no. 9, Apr. 2022, Art. no. 4140, https://doi.org/10.3390/app12094140.

[49] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, Aug. 1998, https://doi.org/10.1109/5.726791.

[50] K. Cho *et al.*, "Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation," in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Doha, Qatar, 2014, pp. 1724–1734, https://doi.org/10.3115/v1/D14-1179.

[51] A. Graves, S. Fernández, and J. Schmidhuber, "Multi-dimensional Recurrent Neural Networks," in *Artificial Neural Networks – ICANN 2007*, Porto, Portugal, 2007, pp. 549–558, https://doi.org/10.1007/978-3-540-74690-4_56.

[52] F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to Forget: Continual Prediction with LSTM," *Neural Computation*, vol. 12, no. 10, pp. 2451–2471, Jul. 2000, https://doi.org/10.1162/089976600300015015.