Journal of Intelligent System and Telecommunications

Journal homepage: https://journal.unesa.ac.id/index.php/jistel/index

A Hybrid CNN-LSTM Approach for Sentiment Analysis of Poverty Issues in East Java

Raissa Alfatikarani[1*],

[1] Depatment of Elctrical Engineering, Faculty of Engineering, Universitas Negeri Surabaya, Suarabaya, Indonesia

E-mail: 24051505001@mhs.unesa.ac.id

ARTICLE INFO

Article History:
Submitted/Received 20 June 2025
First Revised 22 April 2026
Accepted 22 April 2026
First Available Online 29 April 2026
Publication Date 29 April 2026

Keyword:
Sentiment Analysis, CNN-LSTM, Twitter, Poverty, East Java.

ABSTRACT

Often news about poverty in East Java, Indonesia, becomes a topic of conversation on several social media, one of which is Twitter. The purpose of this research is analyzing public sentiment related to poverty by implementing Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM). A dataset of tweets was collected using keyword-based crawling techniques, followed by a series of preprocessing steps including case folding, tokenization, filtering, and labeling. Sentiment classification is categorized as positive and negative using hybrid CNN-LSTM. Evaluation results show that the proposed model achieved a high accuracy of 91%, with equally strong performance in precision, recall, and F1-score. The findings also revealed that "unemployment" was the most dominant topic associated with poverty. This research demonstrates that the hybrid CNN-LSTM architecture is effective for sentiment analysis on short, informal text, and offers valuable insights into public opinion for policymakers.

1. INTRODUCTION

In the current era of digital transformation, discussions that were once confined to traditional media are now widely spread and accessible through online platforms such as Twitter. What may appear as ordinary conversations on social media actually reflects complex public sentiments toward socioeconomic issues, including poverty. With millions of users voicing their opinions online, Twitter has become a valuable source for analyzing how the public perceives various government programs, economic conditions, and social challenges [1].

According to Badan Pusat Statistik (BPS), East Java is one of the provinces with a significant number of people living below the poverty line. In March 2023, it was reported that over 4 million individuals in East Java lived in poverty, with many affected by issues such as unemployment, lack of access to education, and health disparities [2]. Moreover, poverty is often discussed with high intensity on social media, particularly when related to controversial topics like fuel prices, government aid distribution, or social inequality [3]. This makes social media a rich dataset for understanding community perspectives and uncovering patterns in public discourse.

Text-based data such as tweets require excellent techniques to extract important information. Traditional analysis methods feared to be less precise in predicting the contextual and sequential nature of language used in social platforms [4]. Hence, there is a growing interest in leveraging artificial intelligence and deep learning to perform sentiment analysis. Sentiment analysis is a process used to express opinions in text form, especially to determine the author's attitude whether it is positive, negative or neutral [5].

Recent advances in machine learning have enabled more effective sentiment classification. Various algorithms have been employed for this purpose, such as SVM [6], Naive Bayes [7], Logistic Regression [8], and Random Forest [9]. In contrast, deep learning allows neural networks to automatically extract and recognize complex patterns in text data [10], thereby improving performance in tasks such as sentiment classification and opinion mining [11]. One of the commonly used deep learning architectures in sentiment analysis is the Convolutional Neural Network (CNN), which is capable of capturing local semantic patterns such as phrases or n-grams in short text sequences. For instance, studies such as Xinying Chen (2022) demonstrated the effectiveness of CNN in extracting spatial features from short texts [12], while William et al. (2024) showed the strength of LSTM in capturing sentiment flow across long textual sequences [13]. Other works by Swathi P et al. (2023) explored bidirectional LSTM for sentiment classification tasks and achieved promising results [14]. Nevertheless, each model has its limitations when used in isolation. However, CNN is limited in its ability to capture long-range dependencies in sentences, which are often essential for understanding context and sentiment in social media texts. As a result, sentiment analysis using CNN alone often produces moderate accuracy [15].

Other deep learning architectures, such as Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM) networks, have been employed to address this limitation [16]. LSTM is specifically designed to learn temporal dependencies and contextual sequences, making it suitable for tasks involving sequential data like text [17]. Despite its strengths, LSTM can be computationally intensive and slower to train, especially on large datasets or when tasked with extracting spatial features from short, informal sentences such as tweets [18].

In order to overcome the shortcomings of using CNN or LSTM alone, therefore, in this study, the CNN-LSTM hybrid architecture is proposed as a promising solution for performing sentiment analysis on public opinion data related to poverty in East Java. The contribution of this research is that it can inform policymakers' decisions about future policies to reduce poverty in specific areas. Furthermore, it can identify priority areas that require immediate action to improve community well-being.

DOI: http://doi.org/10.26740/jistel.v2n1.p19-28
e- ISSN 3090-613X

2. METHODS

This study follows a structured methodology involving data acquisition, preprocessing, model development, and evaluation. The overall process is designed to ensure effective sentiment classification from Twitter data related to poverty in East Java shown in **Figure 1.**

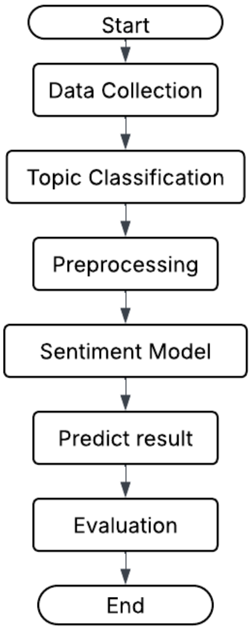

Figure 1. Flowchart Sentiment Analysis Method

2.1. Data Collection

1	id_str	full_text	quote_cou	reply_cour	retweet_c	favorite_c	lang	user_id_st	conversati	username	tweet_url
2	1,7E+18	Beralih ke Cirebon, Kabupaten di timur laut jawa barat ini miliki anggaran berantas kemisk	0	1	0	0	in	1,61E+18	1,7E+18	MasUchiha	https://twit
3	1,7E+18	Banyak keberhasilan Cak @achmadfauzi_wy dalam memimpin Sumenep salah satunya m	0	0	0	2	in	5,47E+08	1,7E+18	hrmnt__	https://twit
4	1,69E+18	Di Jawa Timur, angka pernikahan anak masih tinggi dan justru meningkat selama pandemi.	15	7	200	509	in	1,37E+18	1,69E+18	projectm_	https://twit
5	1,69E+18	1.762 pasangan suami istri di Bojonegoro, Jawa Timur, mengajukan cerai di Pengadilan Ag	0	0	0	0	in	1,32E+08	1,69E+18	officialine	https://twit
6	1,69E+18	Di Tangan Khofifah Indar Parawansa Angka Kemiskinan di Jawa Timur Menurun https://t.	0	0	0	3	in	2,39E+09	1,69E+18	jawapos	https://twit
7	1,69E+18	Jawa Tengah bukan lagi PROVINSI TERMISKIN di Jawa Jateng bahkan lebih baik dari Jawa	38	631	209	1005	in	1,1E+18	1,69E+18	Miduk17	https://twit
8	1,69E+18	Mari tunaikan sedekah untuk mendukung BAZNAS Provinsi Jawa Timur dalam menunjang	0	0	0	2	in	1,69E+18	1,69E+18	baznas_jat	https://twit
9	1,69E+18	Buzzer Ganjen yang sudah tidak relevan Memutar kaset lama untuk rakyat Rakyat Jawa Te	0	0	0	0	in	1,26E+18	1,69E+18	txt_injustic	https://twit
10	1,69E+18	@StefanAntonio__ @Miduk17 @Aryprasetyo85 @03__nakula @ch_chotimah2 @DrPulln	0	0	0	0	in	1,57E+09	1,69E+18	DanishDed	https://twit
11	1,69E+18	Problematika di daerah Madura dr dulu hingga sekarang tak sedikit pun ada yg melek. Hal	0	1	1	2	in	1,63E+18	1,69E+18	bung_mad	https://twit
12	1,69E+18	Problematika di daerah Madura dr dulu hingga sekarang tak sedikit pun ada yg melek. Hal	0	1	0	0	in	1,63E+18	1,69E+18	bung_mad	https://twit
13	1,69E+18	Arahan Prabowo Subianto, Gerindra Komitmen Entaskan Kemiskinan di Jawa Timur #Jokc	0	1	0	0	in	3,67E+09	1,69E+18	AndriK145	https://twit
14	1,69E+18	Per 2023 data dari berbagai media yg menjadikan BPS sebagai rujukan sumber data, angka	0	1	0	0	in	1,34E+18	1,69E+18	GaTypoLa	https://twit
15	1,69E+18	Wakil Gubernur Jawa Timur Emil Dardak mengajak para pemuda untuk bertukar pikiran ur	0	0	0	1	in	1,44E+18	1,69E+18	Andra_cuk	https://twit
16	1,69E+18	Sebanyak 10.228 unit rumah di Kabupaten Ngawi Jawa Timur dikategorikan tak layak huni	0	0	0	1	in	1,36E+18	1,69E+18	CelahId	https://twit
17	1,69E+18	Koperasi menjadi harapan tekan kantong kemiskinan di Jawa Timur â€œEmil Dardakâ€ #	0	0	0	0	in	1,34E+18	1,69E+18	sabdaxkala	https://twit
18	1,69E+18	@Melihat_Indo 5. Kinerja Menurut saya tidak bisa dibandingkan, karena memang tidak da	0	1	0	1	in	3,07E+09	1,69E+18	2020Paner	https://twit
19	1,69E+18	Mari bersedekah untuk mendukung BAZNAS Provinsi Jawa Timur dalam menunjang berbag	0	0	0	1	in	1,69E+18	1,69E+18	baznas_jat	https://twit
20	1,69E+18	Emil Dardak : Koperasi jadi harapan tekan kantong kemiskinan di Jawa Timur. #sigara #BE	0	0	0	0	in	1,44E+18	1,69E+18	_jejakkita	https://twit
21	1,69E+18	Emil Dardak : Koperasi jadi harapan tekan kantong kemiskinan di Jawa Timur #EmilDarda	0	0	0	0	in	1,44E+18	1,69E+18	griyabacak	https://twit
22	1,69E+18	Fraksi Gerindra DPRD Jatim Komitmen Bantu Gubernur Khofifah Entaskan Kemiskinan di Ja	0	0	0	0	in	1,11E+18	1,69E+18	fathw25	https://twit
23	1,68E+18	Workshop Sosial Anggota DPRD Jatim Bahas Kemiskinan Global. ðŸ‘‡ Workshop sosial yai	0	0	0	0	in	21287066	1,68E+18	beritajatim	https://twit
24	1,68E+18	Kali ini pak Muhadjir Effendy turun ke Malang, Jawa Timur melihat masih adanya kemiskin	0	0	1	1	in	1,68E+18	1,68E+18	conggkak	https://twit

Figure 2. Result Of Crawling Dataset

DOI: http://doi.org/10.26740/jistel.v2n1.p19-28
e- ISSN 3090-613X

```
[ ] !pip install pandas
    !curl -sL https://deb.nodesource.com/setup_18.x | sudo -E bash -
    !sudo apt-get install -y nodejs

    Tampilkan output tersembunyi

    # Crawl Data

    filename = 'KEMISKINAN.csv'
    search_keyword = 'KEMISKINAN JAWA TIMUR until:2025-06-15 since:2020-01-03 lang:id'
    limit = 100

    !npx --yes tweet-harvest@latest -o "{filename}" -s "{search_keyword}" -l {limit} --token ""

    Welcome to the Twitter Crawler

    This script uses Chromium Browser to crawl data from Twitter with *your* Twitter auth token.
    Please enter your Twitter auth token when prompted.

    Note: Keep your access token secret! Don't share it with anyone else.
    Note: This script only runs on your local device.
```

Figure 3. Crawling Data Process

Data crawling was employed as the primary technique for acquiring Twitter data. This method involved collecting tweets programmatically using the Twitter API through the Tweepy library, a Python-based wrapper that simplifies interaction with Twitter's RESTful services. The advanced search functionality allowed filtering tweets based on specific keywords such as "kemiskinan," to ensure relevance to the poverty topic under study as shown in Figure 3.

In addition, filters were applied to define the time range or period associated with trending issues, aligning with the temporal scope of the sentiment analysis project. The data collected included tweet content, usernames, timestamps, retweet counts, and other metadata. This approach enabled the construction of a dataset that reflects real-world public discourse, making it suitable for training and evaluating the sentiment classification model [19]. The collected tweets were saved in CSV format and this structured format supports further analysis and model training as shown in Figure 2.

2.2 Preprocessing

Preprocessing steps were essential to clean and normalize the raw tweet data. The following tasks were performed:

- Case folding: All characters in the tweet texts were converted to lowercase [20]. This unification prevents duplication of features that are essentially the same, such as "Poverty" and "poverty".
- Tokenizing: Each tweet was split into individual tokens or words using natural language processing (NLP) libraries. This process is essential for enabling the model to process text at the word level [21].
- Filtering: Removing punctuation, stopwords, and irrelevant symbols.
- Labeling: Each tweet was manually or semi-automatically labeled as Positive, Negative.

2.3 Proposed Architecture

DOI: http://doi.org/10.26740/jistel.v2n1.p19-28
e- ISSN 3090-613X

The architecture depicted in Figure 1 illustrates the proposed CNN-LSTM model used for sentiment classification of tweet data. The model begins with an input layer followed by a series of five consecutive 1D Convolutional (Conv1D) layers. These convolutional layers are responsible for extracting local semantic patterns such as n-grams from the input text data. The use of multiple Conv1D layers allows the model to capture increasingly abstract features. The extracted features are then passed into an LSTM (Long Short-Term Memory) layer with a hidden size of 10 units.

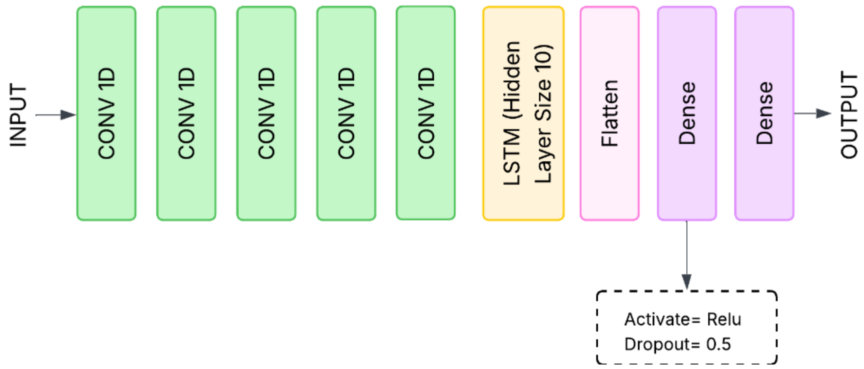

Figure 4. CNN-LSTM Architecture

Batch Size	60
Epoch	10
Conv Filter 1D	Filters=32 , kernel_size=2, padding='same' , activation='relu'
LSTM Size	128, dropout=0.2 , recurrent_dropout=0.2
Layer Dense (1)	64 , dropout=0.5
Layer Dense (2)	16

Figure 5. Parameter of CNN-LSTM Model

The LSTM layer is designed to learn the sequential dependencies and contextual flow across the feature maps produced by the convolutional layers. Following the LSTM, a flatten layer is used to transform the multi-dimensional outputs into a one-dimensional vector. This vector is then forwarded through two fully connected dense layers, which perform the final transformation of learned features into class scores. A Dropout layer with a rate of 0.5 is applied after the first Dense layer to prevent overfitting and enhance generalization. The ReLU (Rectified Linear Unit) activation function is used in the intermediate layers, while the final dense layer uses a softmax function to produce probability distributions over the sentiment classes [22]. This layered design enables the model to leverage the pattern-recognition strength of CNN and the sequential processing capability of LSTM, making it suitable for sentiment analysis on short, noisy, and informal text data like tweets.

3. RESULTS AND DISCUSSION

The learning and testing process has been carried out, the results are explained in the CNN-

DOI: http://doi.org/10.26740/jistel.v2n1.p19-28
e- ISSN 3090-613X

LSTM performance table (**Table 1**), and a comparison has been made with the previous model described in **Table 2**. The illustration of training loss and validation loss is described in **Figure 6**, and the resulting confusion matrix of the proposed system is depicted in **Figure 7**.

Table 1. The Performance of CNN-LSTM

Accuracy	Loss	Precision	Recall	F1-Score
0.91	0.22	0.91	0.91	0.91

Table 1 presents the performance evaluation of the proposed CNN-LSTM hybrid model in terms of accuracy, loss, precision, recall, and F1-score. The model achieved a consistent result of 91% across all evaluation metrics, with a loss value of 0.22. This performance demonstrates that the model is not only capable of learning patterns from training data but also generalizes well to unseen data. The balanced values across precision and recall indicate that the model performs equally well in identifying both positive and negative sentiment [23]. The low loss value further reflects the model's strong convergence during training [24]. These results validate the effectiveness of combining convolutional and recurrent layers, where CNN successfully extracts semantic features from text sequences and LSTM captures temporal dependencies, enhancing the model's contextual understanding of tweet-based sentiment.

Table 2. The Comparison of Previous Model

Model	Accuracy	F1 Score	Precision
SVM [6],	80 %	88 %	71%
Naive Bayes [7]	71%	71 %	71%
Logistic Regression [8]	69 %	69%	69 %
Random Forest [9]	75.29 %	78%	70.22 %
Bi-LSTM [14]	90 %	90 %	90 %
Proposed Hybrid CNN-LSTM	**91%**	**91 %**	**91 %**

Table 2 compares the proposed CNN-LSTM model to several existing machine learning and deep learning approaches previously employed in sentiment classification. Classical machine learning models such as SVM 80% accuracy, 88% F1-score [6] , Naive Bayes 71% [7] , and Logistic Regression 69% [8] exhibit notably lower performance. These models often rely on manually extracted features and cannot fully capture the sequential and nonlinear dependencies present in natural language, especially in informal and noisy social media data [25]. Ensemble learning methods like Random Forest [9] provide slight improvements, with an accuracy of 75.29% and an F1-score of 78%, but still fall short when compared to neural models. The Bi-LSTM model [14], a more advanced deep learning technique, achieves 90% across all metrics, proving the utility of recurrent networks in processing sequential data. However, the proposed hybrid CNN-LSTM model outperforms all previous models, achieving the highest performance with 91% across accuracy, precision, and F1-score. This improvement can be attributed to the synergy between CNN and LSTM components, which allows the model to not only capture local dependencies through convolutional operations but also learn the long-term structure of text sequences via LSTM. Furthermore, unlike standalone CNN models that may struggle with long-distance dependencies or LSTM models that require more computational time, the hybrid architecture achieves a balance between performance and efficiency. This makes it a highly suitable choice for sentiment analysis tasks involving short, dynamic, and high-volume data sources like Twitter. These findings support the proposed method as a scalable solution and robust for understanding public sentiment toward social issues such as poverty in East Java.

DOI: http://doi.org/10.26740/jistel.v2n1.p19-28
e- ISSN 3090-613X

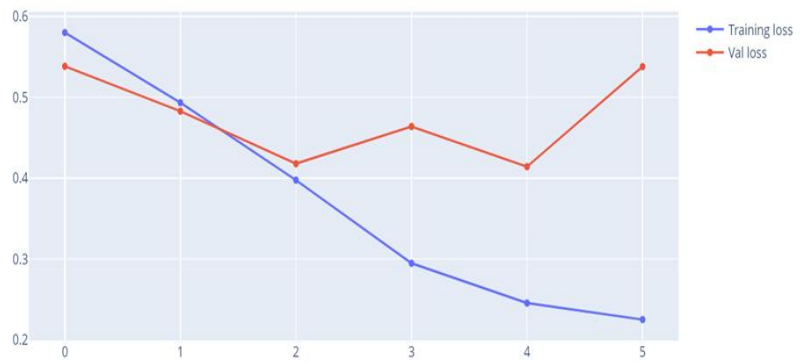

Figure 6. Illustration of Training Loss and Validation Loss

Figure 6 illustrates the trend of training versus validation loss over training process. In this study, the model was initially set to train for 10 epochs. However, Overfitting prevention is achieved by implementing early stopping, which is achieved by incorporating validation loss. Training was halted when no improvement was observed for two consecutive epochs following the lowest validation loss. As a result, the training concluded at epoch 6.

The validation loss serves as a critical indicator of model generalization, which illustrates how well the performance has never been seen before. In this case, it was observed that the validation loss began to fluctuate after the third epoch, indicating potential overfitting if training were to continue. The early stopping mechanism thus ensured the model did not degrade in performance by overtraining on the training set [26]. This strategy supports model robustness and helps preserve predictive performance on new data.

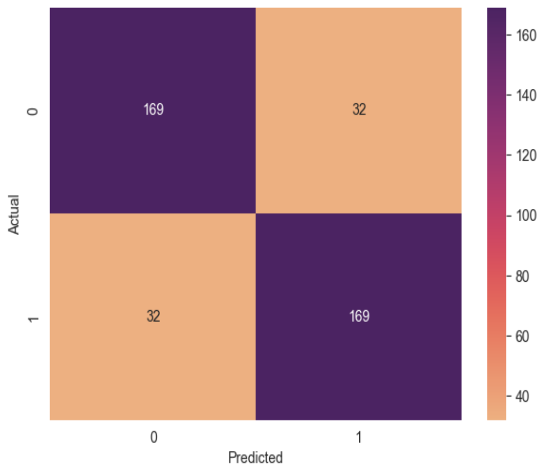

Figure 7. Confusion Matrix

Figure 7 ilustrated the confusion matrix derived from the validation set consisting of 201 data points. The CNN-LSTM model accurately classified 169 samples, while 32 samples were misclassified. The diagonal part of confusion matrix represent correctly predicted sentiments, confirming that the model achieved a high level of classification accuracy. These results support the reliability of the model in identifying sentiment polarity in real-world Twitter data, especially when dealing with class-balanced binary sentiment tasks.

DOI: http://doi.org/10.26740/jistel.v2n1.p19-28
e- ISSN 3090-613X

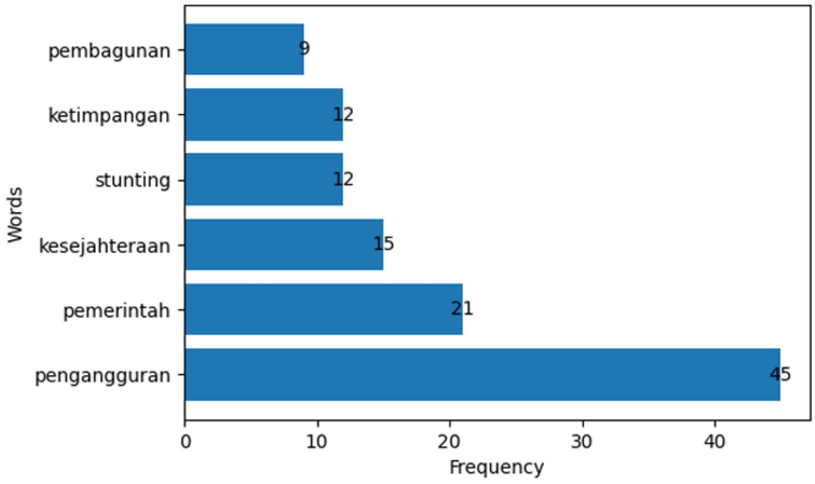

Figure 8. Distribution of Topics Related to the Keyword "Kemiskinan" (Poverty)

The combination of low validation loss, effective early stopping, and strong confusion matrix results demonstrate that the proposed hybrid CNN-LSTM model is both stable and effective for sentiment analysis applications.

Figure 8 presents a horizontal bar chart illustrating the frequency distribution of topics that co-occur with the keyword "kemiskinan" (poverty) in the dataset. The most dominant topic is "pengangguran" (unemployment) with 45 occurrences, indicating strong public association between poverty and joblessness. This is followed by mentions of "pemerintah" (government), "kesejahteraan" (welfare), "stunting", "ketimpangan" (inequality), and "pembangunan" (development). These results highlight the key areas of concern in public discourse surrounding poverty, suggesting that unemployment and government policy are perceived as central issues in poverty alleviation efforts.

4. CONCLUSION

This study applied a hybrid CNN-LSTM model to classify public sentiment on Twitter regarding poverty issues in East Java. The model achieved excellent performance, with 91% accuracy. These results show that combining CNN and LSTM can effectively capture both local word patterns and long-term context in social media text. From the analysis, "pengangguran" (unemployment) appeared as the most frequently mentioned topic related to poverty, followed by government and welfare. This reflects the public's concern toward employment and social support in addressing poverty. Compared to other models such as SVM, Bi-LSTM, and Naive Bayes, the proposed method outperformed them across all evaluation metrics.

For future research, this model can be expanded by incorporating attention mechanisms to improve focus on important words, applying it to multilingual datasets, or developing a real-time sentiment monitoring dashboard for policy evaluation and public feedback analysis.

5. ACKNOWLEDGMENT

Thank you to the Telecommunication Laboratory, Department of Electrical Engineering, Universitas Negeri Surabaya, for facilitating this research.

6. AUTHORS' NOTE

The authors declare that there is no conflict of interest regarding the publication of this article. The authors confirmed that the paper was free of plagiarism.

7. AUTHORS' CONTRIBUTION/ROLE

The authors contributed fully to the entire research process, from data collection and preprocessing to analysis.

8. AI USE AND DECLARATION OF GENERATIVE AI USE

The author uses the Turnitin tool at the end of the finishing paper, this process aims to determine the level of similarity of the paper that has been compiled.

9. REFERENCES

[1] M. Mansoor, "Citizens' trust in government as a function of good governance and government agency's provision of quality information on social media during COVID-19," *Gov. Inf. Q.*, vol. 38, no. 4, p. 101597, 2021, doi: 10.1016/j.giq.2021.101597.

[2] M. A. M. Wibisono and T. Sirait, "Multidimensional Poverty Modeling in Central Java, Di Yogyakarta, and East Java Provinces," *BAREKENG J. Ilmu Mat. dan Terap.*, vol. 18, no. 3, pp. 1939–1954, 2024, doi: 10.30598/barekengvol18iss3pp1939-1954.

[3] G. E. Halkos and P. C. Aslanidis, "Achieving Sustainable Development," 2023.

[4] A. Cas, "A review on sentiment analysis from social media platforms Margarita Rodríguez-Ib a," vol. 223, no. March, 2023, doi: 10.1016/j.eswa.2023.119862.

[5] K. L. Tan, C. P. Lee, and K. M. Lim, "A Survey of Sentiment Analysis: Approaches, Datasets, and Future Research," *Appl. Sci.*, vol. 13, no. 7, 2023, doi: 10.3390/app13074550.

[6] R. Obiedat *et al.*, "Sentiment Analysis of Customers' Reviews Using a Hybrid Evolutionary SVM-Based Approach in an Imbalanced Data Distribution," *IEEE Access*, vol. 10, pp. 22260–22273, 2022, doi: 10.1109/ACCESS.2022.3149482.

[7] C. Villavicencio, J. J. Macrohon, X. A. Inbaraj, J. H. Jeng, and J. G. Hsieh, "Twitter sentiment analysis towards covid-19 vaccines in the Philippines using naïve bayes," *Inf.*, vol. 12, no. 5, 2021, doi: 10.3390/info12050204.

[8] S. Kumar, N. Kaur, Kavita, and A. Joshi, "Tweet Sentiment Analysis using Logistic Regression," *IET Conf. Proc.*, vol. 2023, no. 11, pp. 332–336, 2023, doi: 10.1049/icp.2023.1801.

[9] A. Mardjo and C. Choksuchat, "HyVADRF: Hybrid VADER-Random Forest and GWO for Bitcoin Tweet Sentiment Analysis," *IEEE Access*, vol. 10, no. September, pp. 101889–101897, 2022, doi: 10.1109/ACCESS.2022.3209662.

[10] A. A. Mon and U. N. Surabaya, "Journal of Intelligent System and SURFACE DETECTION FOR QUADRUPED ROBOT USING YOLO-V3 TINY," vol. 1, pp. 13–24, 2024.

[11] S. Sengupta *et al.*, "A review of deep learning with special emphasis on architectures, applications and recent trends," *Knowledge-Based Syst.*, vol. 194, p. 105596, 2020, doi: 10.1016/j.knosys.2020.105596.

[12] X. Chen, P. Cong, and S. Lv, "A Long-Text Classification Method of Chinese News Based on BERT and CNN," *IEEE Access*, vol. 10, pp. 34046–34057, 2022, doi: 10.1109/ACCESS.2022.3162614.

[13] I. O. William and M. Altamimi, "Hierarchical Long Short-Term Memory (LSTM) Model for News Sentiment Analysis," no. June, 2024.

[14] U. B. Mahadevaswamy and P. Swathi, "Sentiment Analysis using Bidirectional LSTM Network," *Procedia Comput. Sci.*, vol. 218, pp. 45–56, 2022, doi:

DOI: http://doi.org/10.26740/jistel.v2n1.p19-28
e- ISSN 3090-613X

10.1016/j.procs.2022.12.400.

[15] J. Lei, H. Wang, Z. Lei, J. Li, and S. Rong, "CNN–Transformer Hybrid Architecture for Underwater Sonar Image Segmentation," *Remote Sens.*, vol. 17, no. 4, pp. 1–32, 2025, doi: 10.3390/rs17040707.

[16] A. C. M. V. Srinivas, C. Satyanarayana, C. Divakar, and K. P. Sirisha, "Sentiment Analysis using Neural Network and LSTM," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 1074, no. 1, p. 012007, 2021, doi: 10.1088/1757-899x/1074/1/012007.

[17] R. K. Behera, M. Jena, S. K. Rath, and S. Misra, "Co-LSTM: Convolutional LSTM model for sentiment analysis in social big data," *Inf. Process. Manag.*, vol. 58, no. 1, 2021, doi: 10.1016/j.ipm.2020.102435.

[18] I. D. Mienye, T. G. Swart, and G. Obaido, "Recurrent Neural Networks: A Comprehensive Review of Architectures, Variants, and Applications," *Information*, vol. 15, no. 9, p. 517, 2024, doi: 10.3390/info15090517.

[19] E. O. Omuya, G. Okeyo, and M. Kimwele, "Sentiment analysis on social media tweets using dimensionality reduction and natural language processing," no. September 2022, pp. 1–14, 2023, doi: 10.1002/eng2.12579.

[20] 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," pp. 35239–35266, 2021.

[21] J. C. Kiamehr *et al.*, "TweetNLP : Cutting-Edge Natural Language Processing for Social Media," no. January 2022, 2020.

[22] A. A. Alkhouly, A. Mohammed, and H. A. Hefny, "Improving the Performance of Deep Neural Networks Using Two Proposed Activation Functions," *IEEE Access*, vol. 9, pp. 82249–82271, 2021, doi: 10.1109/ACCESS.2021.3085855.

[23] M. R. P. Kumar and P. Jayagopal, "Context-sensitive lexicon for imbalanced text sentiment classification using bidirectional LSTM," *J. Intell. Manuf.*, vol. 34, no. 5, pp. 2123–2132, 2023, doi: 10.1007/s10845-021-01866-0.

[24] Y. Wang *et al.*, "F REE M ATCH : S ELF - ADAPTIVE T HRESHOLDING S EMI - SUPERVISED L EARNING," pp. 1–20, 2023.

[25] R. I. Trisanti, A. Nizar, and D. Nur, "Journal of Intelligent System and Predicting Regional Sanitation Conditions Using Support Vector Machines and Neural Networks," vol. 1, pp. 41–50, 2024.

[26] S. H. Al-gburi *et al.*, "EffRes-DrowsyNet : A Novel Hybrid Deep Learning Model Combining EfficientNetB0 and ResNet50 for Driver Drowsiness Detection," pp. 1–36, 2025.