Jurnal Telematika vol. 20 no. 2 2025
e-ISSN: 2579-3772
doi: 10.61769/telematika.v20i2.796

Public Sentiment Toward the Indonesian Capital Relocation Policy on X Using a BiLSTM-CNN Model

Wanda Nugraha[1], Mochamad Tito Julianto[2], Mohamad Khoirun Najib[3], Elis Khatizah[4,*]

[1,2,3,4]*Computational Mathematics Division, School of Data Science, Mathematics, and Informatics, IPB University*
Jln. Meranti Kampus IPB Dramaga, Bogor, Indonesia

wnddanugraha@apps.ipb.ac.id
mtjulianto@apps.ipb.ac.id
mkhoirun@apps.ipb.ac.id

*Correspondence: elis_khatizah@apps.ipb.ac.id

Abstract*— The development of Indonesia's new capital city, Ibu Kota Nusantara (IKN), is an innovative government policy that has sparked diverse public responses. This study aims to explore sentiment trends on the social media platform X to understand public perceptions of the policy. Additionally, a sentiment classification model combining Bidirectional Long Short-Term Memory (BiLSTM) and Convolutional Neural Network (CNN) was developed and optimized through hyperparameter tuning. Exploratory analysis showed that positive sentiment dominated at 46%, followed by negative at 30% and neutral at 24%. The classification model achieved a test accuracy of 78% and an average accuracy of 81% across 10-fold cross-validation, with a standard deviation of 0.006. The achieved accuracy, together with the low cross-validation standard deviation, indicates that the BiLSTM-CNN model demonstrates stable and reliable performance.

***Keywords*— Ibu Kota Nusantara (IKN), public sentiment, X social media, BiLSTM-CNN, sentiment analysis**

***Abstrak*— Pembangunan ibu kota baru Indonesia, Ibu Kota Nusantara (IKN), merupakan kebijakan pemerintah yang inovatif dan telah memicu respons publik yang beragam. Studi ini bertujuan untuk menganalisis tren sentimen di *platform* media sosial X guna memahami persepsi publik terhadap kebijakan tersebut. Selain itu, model klasifikasi sentimen yang menggabungkan *bidirectional long short-term memory* (BiLSTM) dan *convolutional neural network* (CNN) dikembangkan dan dioptimalkan melalui penyesuaian hiperparameter. Analisis eksploratori menunjukkan bahwa sentimen positif mendominasi 46%, diikuti oleh sentimen negatif 30% dan netral 24%. Model klasifikasi mencapai akurasi uji sebesar 78% dan akurasi rata-rata 81% pada *10-fold cross-validation*, dengan simpangan baku 0,006. Akurasi yang dicapai, bersama dengan simpangan baku *cross-validation* yang rendah, menunjukkan bahwa model BiLSTM-CNN menunjukkan kinerja yang stabil dan andal.**

***Kata Kunci*— Ibu Kota Nusantara (IKN), sentimen publik, media sosial X, BiLSTM-CNN, analisis sentimen**

I. Introduction

In Indonesia, social media has become one of the main platforms for people to share their opinions on many issues, including government policies [1]. One major policy that has received significant public attention is the plan to build a new national capital in East Kalimantan. Although this plan is not new, it remains important because it affects the interests of the public, the government, and other stakeholders. The development of the Nusantara Capital City, or Ibu Kota Nusantara (IKN), aims to move the focus of national development from Java Island to other regions, and to create new economic opportunities in East Kalimantan. With an estimated budget of IDR 466 trillion [2] and strong efforts to promote its benefits, this policy has received mixed reactions from the public [3]. Some people support the plan, hoping it will boost the economy, improve infrastructure and technology, and help address problems such as overpopulation, traffic jams, and environmental damage in Jakarta [4]. However, others are critical of the project, mainly because of its high cost, possible social and political effects, environmental concerns from land clearing, and doubts about its long-term success [5].

The diverse views on the capital relocation highlight the need for further efforts to understand public perceptions of this policy. Understanding public responses can provide the government with valuable insights to design strategic actions that address public aspirations and concerns [6]. In the field of text mining, sentiment analysis, part of Natural Language Processing (NLP), can be used for this purpose. Furthermore, sentiment analysis using deep learning can more deeply explore people's emotions, opinions, and attitudes [7]. This is especially important for processing data from digital interactions on social media, which often contain informal language, abbreviations, emojis, and sarcasm, elements that traditional methods find difficult to handle.

However, without a systematic and reliable analysis of public sentiment, policymakers may misunderstand the level of public support or opposition to the IKN relocation policy. This misunderstanding can lead to ineffective communication strategies and may reduce public trust. Therefore, a sentiment analysis approach that is both methodologically sound and computationally practical is needed to better capture nuanced public opinions expressed on social media.

One of the commonly used deep learning architectures for text analysis is the Recurrent Neural Network (RNN), which is designed to retain information from previous steps in a data sequence, allowing the model to recognize patterns more accurately in sequential data [8]. Bidirectional Long Short-Term Memory (BiLSTM) is a deep learning technique derived from the RNN architecture, specifically designed to address classification problems involving sequential data. BiLSTM allows processing data sequences in two directions, forward and backward, enabling the model to capture a more complete context from earlier and later information within a text sequence [9].

Previous studies have shown the effectiveness of BiLSTM in various sentiment analysis contexts, such as when compared to LSTM and Logistic Regression on Shopee e-commerce reviews [10], LSTM and GRU for emotion classification [11], BERT and ensemble methods for emotion recognition in Indonesian product reviews [12], as well as SVM, Random Forest, CNN, and LSTM for evaluating public service perceptions [13]. The exclusive use of BiLSTM has also been tested on Twitter reviews with an accuracy of 58.56% [14], identification of bullying patterns on social media with 82.04% accuracy [15], multi-label hate speech classification with 82.31% accuracy [16], and analysis of government policy during the COVID-19 pandemic with an F1-score of 76.67% [17]. Public sentiment analysis of service platforms such as Grab Indonesia [18] and MyPertamina [19] also demonstrated the strong performance of BiLSTM. Beyond the human language, this method has also been effectively applied in the analysis of biological sequences [20], [21], [22], treating the sequence as the language of life.

Nevertheless, the application of BiLSTM in sentiment analysis related to public policy, particularly the development of IKN, remains limited. Using a document classification approach on data collected from social media and news media around May 2022, one study [23] found that negative public sentiment toward the policy was relatively more dominant. In addition, basic machine learning methods such as the Naive Bayes Classifier achieved 76.74% accuracy [24], while K-Nearest Neighbor and Random Forest showed lower accuracies of 58.25% and 45.05%, respectively [25], and SVM achieved more than 90% accuracy [26]. Existing deep learning approaches, which generally perform better on larger datasets, have employed CNN with 66.8% accuracy [27], and IndoBERT, which outperformed simpler machine learning methods with 83% accuracy [28]. Across these studies, sentiment labeling was conducted using different approaches: manual labeling [24], [25], a combination of manual annotation and LLM-assisted labeling [28], and fully automated lexicon-based labeling tools [26]. While transformer-based models such as IndoBERT demonstrate strong performance, they typically require significant computational resources for pre-training and fine-tuning [29], which may not be feasible for smaller-scale academic research projects.

Recent studies have shown that, without relying on a transformer, combining CNN and RNN architectures, including LSTM and BiLSTM, can yield superior performance compared to using them individually. CNNs are effective in extracting local patterns or spatial features, while LSTMs excel at capturing long-term dependencies in sequential data [30]. Various combinations, such as CNN-BiLSTM [31], CNN-RNN [32], and CNN-LSTM have been shown to improve model accuracy and efficiency for dimensional sentiment analysis [33]. Similar results were also reported in studies integrating BiLSTM and CNN [30], [34]. Therefore, using a BiLSTM-CNN hybrid is justified, as it captures both local patterns and long-term dependencies, which are essential for understanding nuanced sentiment in social media text [35]. However, no existing research has explicitly applied the BiLSTM-CNN architecture to evaluate public opinion on Indonesia's new capital city relocation policy, IKN. To address this gap, this study adopts a BiLSTM-CNN approach for sentiment analysis on the IKN issue and evaluates its performance.

II. METHODOLOGY

The data used in this study are secondary, consisting of posts from the social media platform X (formerly Twitter). The data was collected automatically using the command-line tool Twitter Harvest v2.6.1, which utilizes an access token and specific search parameters. In this study, Twitter Harvest was run using the researcher's personal access token, with the following parameters: the keyword "Ibu Kota Nusantara", the date range from January 8 to January 26, 2022, and the language set to Indonesian. This period was chosen to capture the dynamics of public sentiment before and after the ratification of the National Capital City Law (UU IKN) by the Indonesian House of Representatives (DPR RI) on January 18, 2022. During this time, media coverage and public statements from government officials regarding the IKN relocation policy were highly intense, which could influence opinions expressed by social media users on platform X. The data collection process was conducted using the Python programming language and executed on Google Colab. With these parameters, a total of 10244 posts were successfully gathered. It should be noted that the dataset size was limited by the approximately 600 posts per day retrieval constraint, which may affect the representativeness of the collected data and the generalizability of the analysis.

In general, this study consists of seven main steps: data preprocessing, data exploration, data splitting, feature engineering, hyperparameter tuning, model evaluation, and validation using k-fold cross-validation. We used an automated lexicon-based tool for sentiment labeling, which demonstrated excellent model performance in [26], despite earlier findings indicating that such methods achieve only around 65% accuracy. This automated labeling approach was also chosen to minimize human intervention, thereby reducing the time and cost of labeling large volumes of social media data.

Data preprocessing is the first stage in preparing data for analysis and involves several cleaning and transformation steps. Irrelevant elements such as mentions, links, numbers, and emoticons are removed, and all text is converted to lowercase through case folding. Normalization follows, replacing abbreviations and correcting spelling based on the Colloquial Indonesian Lexicon. Sentiment labelling is then performed automatically using the *valence aware dictionary for sentiment reasoning* (Vader) method, which requires translating the text into English beforehand. The processed text is subsequently tokenized, stopwords are removed, and words are reduced to their root forms through stemming. While this approach allows efficient labeling, it may introduce bias due to translation errors and the limited ability of a general English lexicon to capture colloquial Indonesian expressions, sarcasm, and context-specific sentiment. Therefore, the resulting sentiment labels should be interpreted as approximate representations of public sentiment rather than exact ground truth.

Data exploration is conducted to understand the dataset's characteristics, with a particular focus on the distribution of sentiment classes. This distribution is typically visualized using pie charts. To address class imbalance, data augmentation techniques are employed, specifically back translation and synonym replacement. Back translation entails translating the original Indonesian text into English and then back into Indonesian using Google Sheets, effectively increasing the dataset size. Synonym replacement involves substituting certain words with their synonyms from an Indonesian thesaurus, implemented in Python, to balance the number of samples across minority classes to match those in the majority class.

The preprocessed data was then split into training (70%), validation (15%), and test (15%) sets, following standard machine learning practice. The validation set helps prevent overfitting and is used during hyperparameter tuning to optimize model performance before testing on unseen data. In the feature engineering stage, categorical labels were converted to numerical representations using label encoding and one-hot encoding, and sequences were padded to a maximum length of 51 tokens by adding zero tokens to the ends of shorter sequences. Word embeddings were used as the model's embedding layer with 64 dimensions.

The sentiment classification model uses a BiLSTM-CNN architecture, starting with a BiLSTM layer followed by a one-dimensional convolutional layer (see Figure 1). The model is trained for 30 epochs with a batch size of 32 to accommodate computational resource limitations. At each epoch, the model is trained on the training data and validated to monitor performance. The model with the best validation result is saved and tested on the test data. Optimization uses the Adam optimizer with categorical cross-entropy loss, and early stopping with a patience of three prevents overfitting by halting training if no improvement is observed over three consecutive epochs. At each epoch, the model is trained on the training data and validated to monitor performance. The model with the best validation result is saved and tested on the test data. Optimization uses the Adam optimizer with categorical cross-entropy loss, and early stopping with a patience of three prevents overfitting by halting training if no improvement is observed over three consecutive epochs.

To build an optimal model architecture, hyperparameter tuning was performed using Bayesian Optimization. The search and validation process involved building and training models with various hyperparameter configurations. The hyperparameter search space is detailed in Table I, with 30 trials conducted to find the optimal settings. The selection of these hyperparameters is based on their critical role in influencing model performance. Referring to Figure 1 and Table I, hyperparameters such as dropout rate (n_1), recurrent dropout rate (n_4), and LSTM dropout rate (n_3) was used to prevent overfitting and improve model generalization. The number of LSTM units (n_2), dense layer units (n_6), and Conv1D filters (n_5) were tuned to adjust the model's capacity to extract and process information from textual data. Meanwhile, the learning rate was tuned as it is a key factor determining the efficiency and stability of the training process.

The results of the best-trained model will be re-evaluated using a confusion matrix and several key evaluation metrics, including accuracy, precision, recall, and f1-score. To ensure the model is more stable and accurate, it will also be evaluated using 10-fold cross-validation, and the standard deviation will be calculated to observe how consistent the accuracy is across each fold. A summary of the stages carried out in this study is presented in Figure 2.

To provide a clear overview of both the limitations and capabilities of the system used throughout the experimental procedures in this study, the specifications of the hardware and software employed are outlined as follows. The research was conducted using a laptop equipped with a 1.4 GHz Quad-Core Intel Core i5 processor, an Intel Iris Plus Graphics 645 GPU with 1536 MB memory, 8 GB RAM, and a 128 GB SSD storage drive. The operating system used was MacOS Big Sur version 11.2.1. Supporting software included the Python programming language version 3.10.9, executed through the Google Collaboratory platform, along with auxiliary tools such as Microsoft Excel 365 and Google Chrome browser version 135.0.7049.85 (Official Build) (x86_64). In addition, several Python packages were utilized in the implementation process, including numpy 1.26.4, pandas 2.2.3, nltk 3.9.1, matplotlib 3.10.0, seaborn 0.13.2, re 2.2.1, sastrawi 1.0.1, deep-translator 1.11.4, vaderSentiment 3.3.2, tensorflow 2.16.2, tensorflow keras 3.8.0, and scikit-learn 1.6.1.

The results of the data exploration to understand the distribution of sentiment class proportions from the Vader automatic labeling are presented in a pie chart in Figure 3. The chart shows that the majority of text posts related to the IKN policy are categorized as having positive sentiment, accounting for 46%. Meanwhile, posts with negative sentiment comprise 30%, and 24% of the posts fall into the neutral sentiment category.

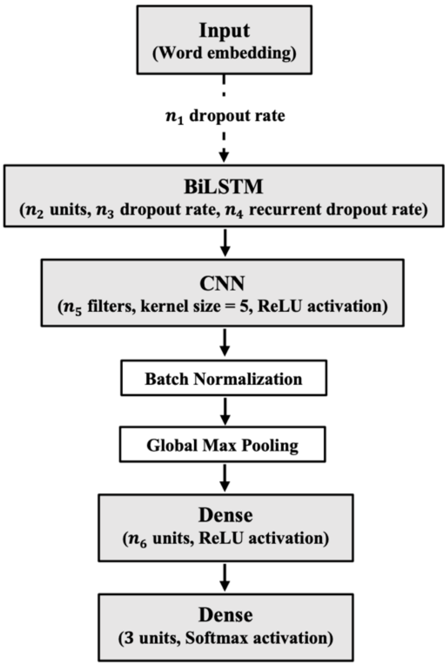

Figure 1 A proposed BiLSTM-CNN model architecture

TABLE I
THE HYPERPARAMETER SEARCH SPACE IN THE TUNING PROCESS

Hyperparameter	**Range of Values**	**Stepsize**
Dropout rate	0.3 – 0.6	0.05
LSTM units	64 – 256	32
Dropout LSTM	0.3 – 0.7	0.05
Reccurent dropout	0.3 – 0.7	0.05
Conv1D filters	32 – 256	32
Dense units	16 – 128	16
Learning rate	1e-5 – 1e-3	log

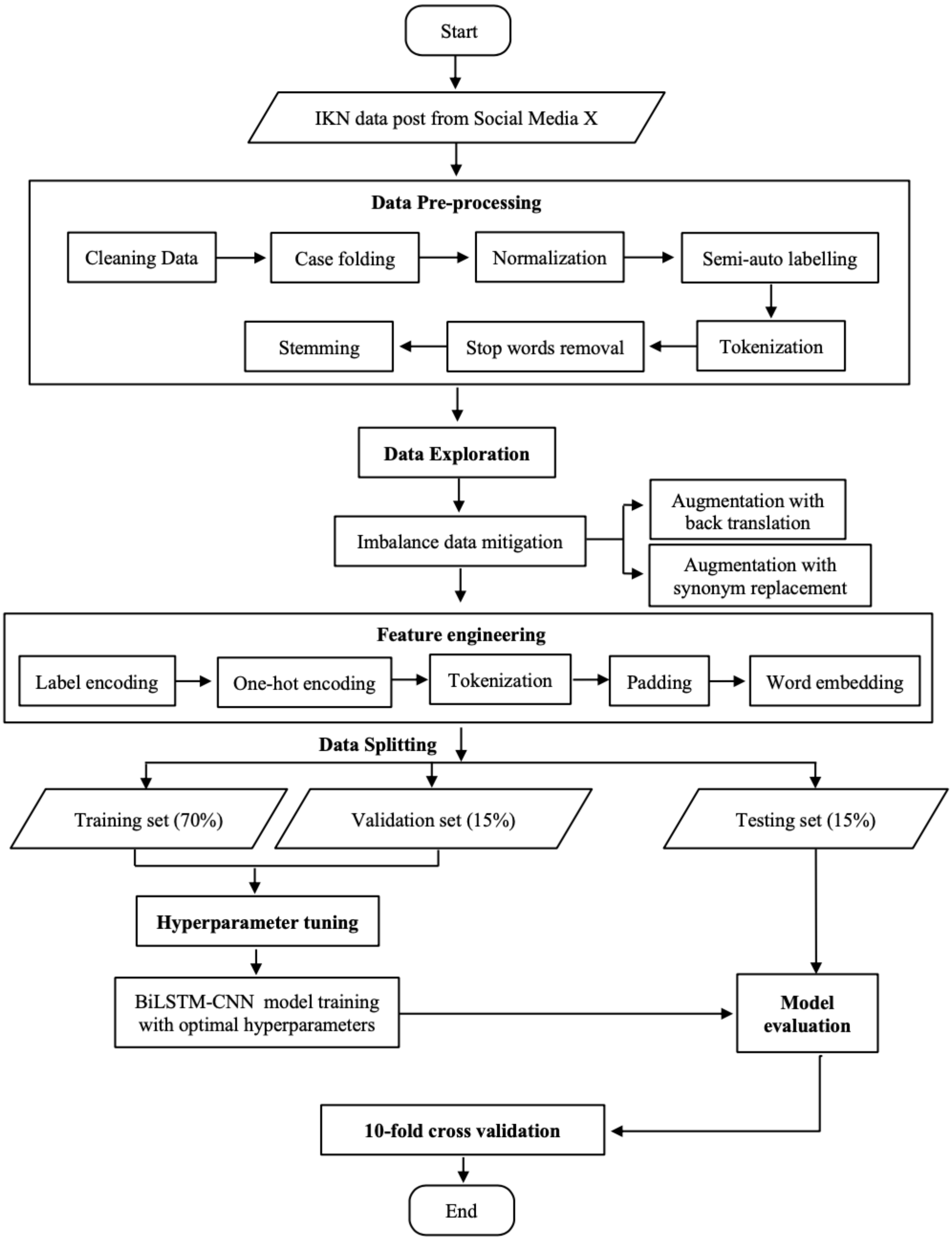

Figure 2 The methodological framework

III. Results and Discussion

In general, these findings indicate that during the period before and after the ratification of the IKN development policy, specifically from January 8 to 26, 2022, users on platform X tended to express positive sentiment toward the policy. This trend is in direct contrast to previous findings reported in studies [23], [24], [25], [27], [28], which reported a higher proportion of negative sentiment in IKN-related posts. For example, [28] analyzed data from April 2017 to April 2025 and found more negative than positive statements. The dominance of negative sentiment was also observed in studies based on datasets collected from January to June 2020 [24], from February 1 to May 30, 2020 [27], and from June to September 2023 [25].

The positive dominance in our dataset may be influenced by high media coverage and frequent government statements during the ratification period, which likely framed the policy in a favorable light. Intense reporting and official communications around the date could have shaped public perceptions and encouraged positive expressions on social media [36]. Additionally, our findings align with the most recent study [26], which collected Indonesian and English tweets from January 2019 to January 2025, showing that positive sentiment was dominant in both languages. Furthermore, the study in [26] labeled Indonesian tweets using the Inset Lexicon, a specifically designed Indonesian lexicon, whereas English tweets were labeled using the Vader lexicon, a commonly used English lexicon. Thus, our results contribute an alternative perspective by labeling Indonesian tweets through translation and using an English-based tool. Examples of posts for each sentiment category are provided in Tables II, III, and IV.

Figure 3 reveals a significant class imbalance, with the positive sentiment class dominating the dataset, while the negative and neutral classes are represented by substantially fewer samples. This imbalance can adversely affect classification performance, as models tend to become biased toward the majority class, making it more difficult to accurately learn and predict patterns in the minority classes. To address this issue, a text data augmentation process was implemented to balance the distribution of sentiment classes.

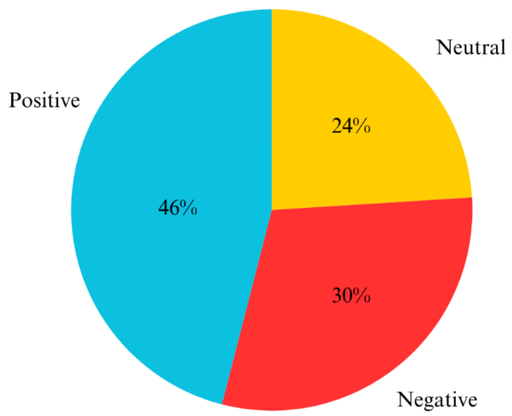

Figure 3 The proportion of sentiment class from the Vader automatic labeling

TABLE II
EXAMPLES OF TEXT POSTS FOR POSITIVE SENTIMENT

No.	Post Text	Label
1	Yg demo paling PKS cabang kalimantan timur tuh.. Saya yakin mayoritas masyarakat kalimantan timur antusias Ibu Kota Nusantara ada di kalimantan timur	positive
2	Alhamdulillah undang undang Ibu Kota Nusantara Sah Ibu Kota RI Bukan Jakarta Lagi 1 Bulan Lagi Jokowi Angkat Kepala Otorita Ibu Kota nusantara https://t.co/48L9d6JmOi	positive
3	@KompasTV Semoga Ibu Kota Nusantara menjadi simbol kemajuan Indonesia lanjutkan dan sukseskan Rakyat Indonesia Mendukung Penuh untuk keberhasilan Ibu Kota Nusantara.	positive

The goal was to increase the number of negative and neutral samples to match the number of positive samples. Two augmentation techniques, back translation and synonym replacement, were employed, and the example results are presented in Table V.

After applying data augmentation with back-translation and synonym replacement, the previously imbalanced dataset became balanced, with a significant increase in the number of samples. The total number of samples rose to 25818, with an equal distribution across sentiment classes, each consisting of 8606 statements.

TABLE III
EXAMPLES OF TEXT POSTS FOR NEGATIVE SENTIMENT

No.	Post Text	Label
1	#TolakUUIbu Kota Nusantara Para cukong dpt proyek besar2an pembangunan Ibu Kota Nusantara tak bermartabat bersumber dr utang yg menggila bikin cilaka generasi kedepan	negative
2	Alasan yg dipakai presiden itupun sangat rasional . Akan lebih fokus pada peningkatan kesejah teraan rakyatnya. Dan pada akhirnya legacy @jokowi yg dikenang rakyat adl bkn Ibu Kota Nusantara kereta cepat atau jln tol . Tp utang yg membebani pemerintahan selanjutnya. MASYAALLAH..	negative
3	Kilurah sebenarnya berat melakukan karena desakan para cukong OLIGARKI kilurahpun tunduk dan harus segera membangun.tidak adanya urgensi pemindahan Ibu Kota Nusantara semua pesanan.	negative

TABLE IV
EXAMPLES OF TEXT POSTS FOR NEUTRAL SENTIMENT

No.	Post Text	Label
1	Menteri PUPR Ungkap Desain Istana di Ibu Kota Nusantara Baru Sudah Ada Pembangunan Tunggu Instruksi Presiden #Nasional https://t.co/qhnXmvYyDh https://t.co/kgtU43xQlS	neutral
2	@conan_idn Sampe sekarang belum paham kenapa harus pindah Ibu Kota Nusantara..urgensjnya apa y..?	neutral
3	Menurut Sri Mulyani jika pembangunan Ibu Kota Nusantara tak bisa menggunakan dana PEN maka nantinya anggaran akan diambil dari pos anggaran Kementerian PUPR https://t.co/iq9eB0yMhP	neutral

TABLE V
THE EXAMPLE RESULTS OF BACK TRANSLATION AND SYNONYM REPLACEMENT

No.	Stemming output	Translation output	Back translation output	Synonym replacement output
1	kerja rumah badan otorita kota nusantara	*homework for nusantara city authority bodies*	pekerjaan rumah bagi otoritas kota nusantara	operasi bangunan parlemen otorita kota nusantara
2	rencana kota nusantara kritik sesuai tata urut ilmu	*Nusantara city plans are criticized according to the order of science*	Rencana kota nusantara dikritik sesuai tatanan ilmu pengetahuan	rencana praja nusantara iktirad sesuai sistem urut kepandaian

The next step involved hyperparameter tuning to optimize the BiLSTM-CNN architecture's performance. Bayesian optimization was employed to identify the optimal hyperparameter combination within the defined search space, as detailed in Table I. A total of 30 hyperparameter search trials were conducted. The results of these trials are presented in the validation accuracy scores for each trial as shown in Figure 4. The highest validation accuracy was achieved in Trial ID 13, with a score of 0.79. In this trial, the model used a dropout rate of 0.40, 96 LSTM units, an LSTM dropout rate of 0.60, a recurrent dropout rate of 0.45, 64 Conv1D filters, 32 dense units, and a learning rate of 0.000416. Trial ID 13 achieved the highest validation accuracy, making it the optimal hyperparameter combination for the model.

To evaluate the model's performance, a confusion matrix is used alongside metrics such as accuracy, precision, recall, and F1-score. The confusion matrix of the model after hyperparameter tuning is shown in Figure 5, and the detailed evaluation metrics are presented in Table VI.

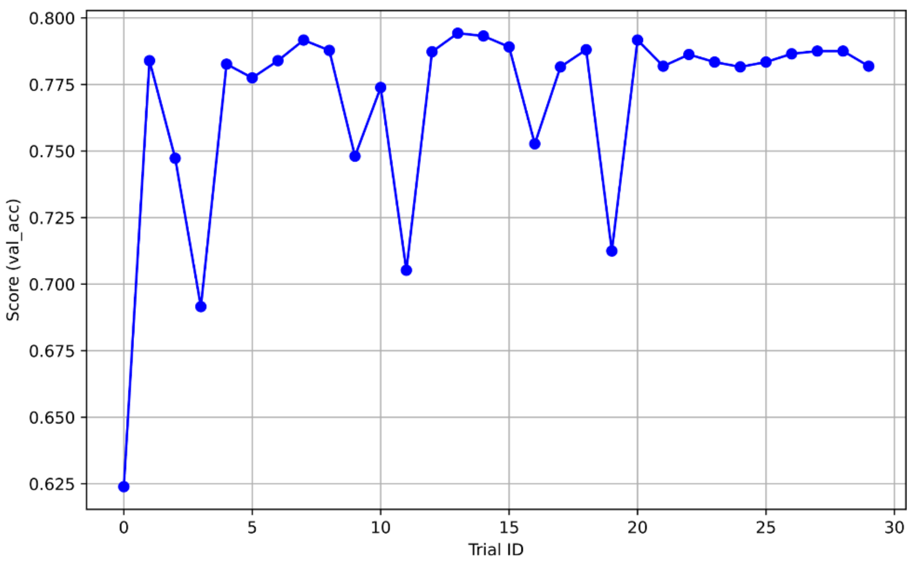

Figure 4 The progress of the tuning process among all trials

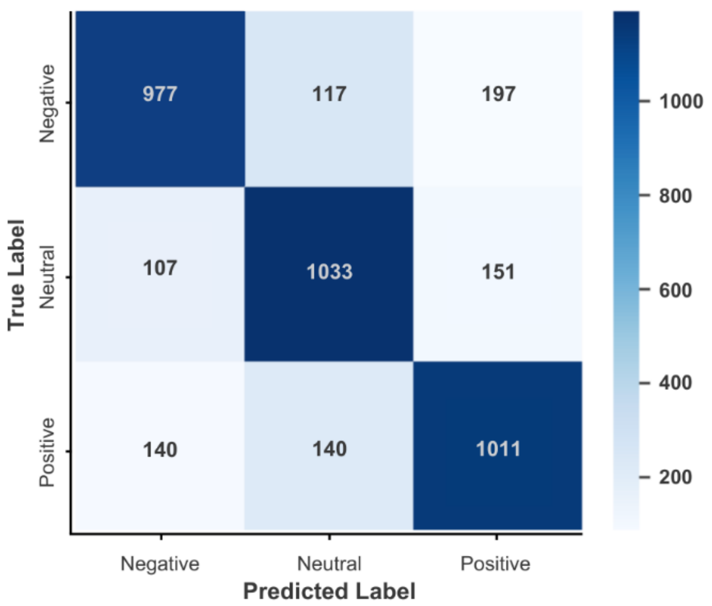

Figure 5 Confusion matrix for the proposed BiLSTM-CNN model

Based on Table VI, the model achieved an accuracy of 0.78, indicating reliable performance across all sentiment classes. Class-wise metrics are relatively balanced, with neutral and negative showing slightly higher precision, recall, and F1-scores (0.80) compared to Positive (0.74–0.78), indicating that the model performs slightly better at identifying negative and neutral sentiments than positive ones. Examination of the confusion matrix reveals specific patterns of misclassification: negative posts were occasionally misclassified as positive (197) or neutral (117); neutral posts were occasionally misclassified as positive (151) or negative (107); and positive posts were evenly misclassified as negative or neutral (140 each). Although there is a slight tendency for negative and neutral posts to be misclassified as positive, these errors are minor relative to the number of correctly predicted instances and do not substantially affect the class-wise metrics.

These observations indicate that the BiLSTM-CNN model performs well, particularly in detecting neutral and negative sentiments. The main limitation lies in distinguishing positive sentiment from other classes in ambiguous or mixed-expression posts. Interpreting these misclassifications highlights areas for potential improvement, such as ensuring realistic data balancing during preprocessing.

To further validate the model’s robustness, 10-fold cross-validation was conducted to ensure that performance was not dependent on a specific data split. The detailed results of the 10-fold cross-validation are presented in Table VII. Based on these results, the model achieved an average accuracy of approximately 0.81 with a low standard deviation of 0.00614. This relatively high accuracy and small variation indicate that the model is stable, consistent, and generalizes well across different subsets of the data, complementing the evaluation based on the single test split.

Overall, the main results demonstrate a systematic approach to developing a reliable sentiment classification model for the IKN policy discourse. Starting with an imbalanced sentiment class distribution, the data were balanced through augmentation techniques to ensure fair representation across classes. The

TABLE VI
PERFORMANCE METRICS FOR EACH SENTIMENT CLASS

Sentiment class	**Performance metrics**			
	Accuracy	**Precision**	**Recall**	**F1-score**
Positive		0.74	0.78	0.76
Negative	0.78	0.80	0.76	0.78
Neutral		0.80	0.80	0.80

TABLE VII
10-FOLD CROSS VALIDATION RESULTS

No	**Fold**	**Accuracy**
1	Fold 1	0.820294
2	Fold 2	0.802479
3	Fold 3	0.810225
4	Fold 4	0.809450
5	Fold 5	0.807901
6	Fold 6	0.805964
7	Fold 7	0.822618
8	Fold 8	0.812548
9	Fold 9	0.806277
10	Fold 10	0.815963
Mean		0.811372
Standard deviation		0.006164

model's performance was further optimized via Bayesian hyperparameter tuning, yielding a configuration that achieved strong accuracy and stability. Evaluation through standard metrics and reinforced by 10-fold cross-validation confirmed the robustness of the BiLSTM-CNN model. While direct comparison is not possible due to differences in datasets, the accuracy achieved in this study appears notably high when viewed alongside other works addressing the same task of classifying IKN sentiment in [24], [25], [27], underscoring the effectiveness of the methods applied.

IV. CONCLUSION

This study demonstrates that the BiLSTM-CNN model can effectively classify public sentiment toward the IKN relocation policy on the social media platform X. Data exploration revealed that positive sentiment dominates at 46%, followed by negative at 30% and neutral at 24%. To address class imbalance, data augmentation techniques, including back translation and synonym replacement, were applied, resulting in a balanced dataset with 8,606 samples per sentiment class. The BiLSTM-CNN model was optimized via Bayesian hyperparameter tuning, yielding an optimal configuration of 96 LSTM units, dropout of 0.40, LSTM dropout of 0.60, recurrent dropout of 0.45, 64 Conv1D filters, 32 dense units, and a learning rate of 0.000416. The model achieved 78% accuracy on the test set and an average accuracy of 81% with a low standard deviation of 0.006 in 10-fold cross-validation, indicating stable and consistent performance across all sentiment classes.

While these results are encouraging, certain limitations should be noted to guide future improvements. The data labeling process relied on an automated method using Vader, which, without expert manual verification, may have led to misclassifications and reduced objectivity. In addition, data collection was limited by the use of an unverified personal account, restricting retrieval to approximately 600 posts per day [37], thereby limiting the dataset size. Addressing these challenges in future research through expert-assisted manual annotation, the involvement of representative annotators, and the addition of the dataset has the potential to further enhance the model's performance.

REFERENCES

[1] G. K. Nathanael, D. Angeliawati, and M. Muhadjir, "The power of social media to influence people in Indonesia: Case study of crazy rich scam and big data of election postponement," *JPPI (Jurnal Penelitian Pendidikan Indonesia)*, vol. 9, no. 1, Mar. 2023, doi: 10.29210/020221755.

[2] K. S. Pribadi and T.-K. Chan, "New capital city," in *Construction in Indonesia*. London, UK: Routledge, 2022, pp. 109–118, doi: 10.1201/9781003149866-7.

[3] M. U. Ikfina, "Indonesia's national capital relocation policy: An analysis of political support in parliament and the public responses on Twitter," *Journal of Governance and Public Policy*, vol. 11, no. 1, pp. 1–13, Feb. 2024, doi: 10.18196/jgpp.v11i1.17760.

[4] Y. Yohanes, S. I. Mayrisa, V. Tiyana, A. P. Indriyanti, P. N. Lestari, and M. N. Fatihah, "Analisis potensi wilayah Provinsi Kalimantan Timur terhadap pembangunan Ibu Kota Nusantara (IKN)," *Kultura: Jurnal Ilmu Hukum, Sosial, dan Humaniora*, vol. 2, no. 5, pp. 129–134, 2024.

[5] A. Aziz SR, "Pemindahan ibu kota negara perspektif civil society," *Journal of Politics and Policy*, vol. 5, no. 2, pp. 104–118, Dec. 2023, doi: 10.21776/ub.jppol.2023.005.02.03.

[6] A. Kharisma and I. Ernawati, "Sentimen analisis opini masyarakat Jakarta pada kinerja pemerintah Jakarta terhadap isu tenggelamnya Jakarta menggunakan algoritma support vector machine," in *Seminar Nasional Mahasiswa Ilmu Komputer dan Aplikasinya (SENAMIKA)*, 2023, pp. 488–497.

[7] J. Husna and M. D. A. Widirahayu, "Pemodelan topik dan analisis sentimen pada 'Voices of History: 50 Iconic Speeches' menggunakan pendekatan natural language processing," *Jurnal Ilmiah Manajemen Informasi dan Komunikasi*, vol. 8, no. 1, pp. 15–24, Jul. 2024, doi: 10.56873/jimik.v8i1.330.

[8] S. Das, A. Tariq, T. Santos, S. S. Kantareddy, and I. Banerjee, "Recurrent neural networks (RNNs): Architectures, training tricks, and introduction to influential research," in *Recurrent Neural Networks*. Cham, Switzerland: Springer, 2023, pp. 117–138, doi: 10.1007/978-1-0716-3195-9_4.

[9] J. Odii, C. Ofoegbu, C. L. Okpalla, and O. I. Christopher, "Emotion detection using a bidirectional long short-term memory (BiLSTM) neural network," *International Journal of Research and Publication Reviews*, 2023. [Online]. Available: www.ijrpr.com

[10] D. Khoirunnisa and E. W. Pamungkas, "Sentiment analysis on Shopee app user feedback: A comparison of LSTM and BiLSTM algorithms," in *2025 International Conference on Smart Computing, IoT and Machine Learning (SIML)*, IEEE, Jun. 2025, pp. 1–5, doi: 10.1109/SIML65326.2025.11081164.

[11] A. Glenn, P. LaCasse, and B. Cox, "Emotion classification of Indonesian tweets using bidirectional LSTM," *Neural Computing and Applications*, vol. 35, no. 13, pp. 9567–9578, May 2023, doi: 10.1007/s00521-022-08186-1.

[12] R. Pramana, M. Jonathan, H. S. Yani, and R. Sutoyo, "A comparison of BiLSTM, BERT, and ensemble method for emotion recognition on Indonesian product reviews," *Procedia Computer Science*, vol. 245, pp. 399–408, 2024, doi: 10.1016/j.procs.2024.10.266.

[13] D. U. Iswavigra, "Sentiment analysis to evaluate public service perception among Surakarta City residents using the BiLSTM model," *Journal of Informatics and Telecommunication Engineering*, vol. 9, no. 1, pp. 229–239, 2025, doi: 10.31289/jite.v9i1.15498.

[14] R. W. Pratiwi, S. F. Handayani, D. Dairoh, and D. I. Af'idah, "Sentiment analysis on Twitter reviews data using the bidirectional long short-term memory (BiLSTM)," in *AIP Conference Proceedings*, 2024, Art. no. 030021, doi: 10.1063/5.0199238.

[15] A. I. Safitri and T. B. Sasongko, "Sentiment analysis of cyberbullying using bidirectional long short-term memory algorithm on Twitter," *Jurnal Teknik Informatika (JUTIF)*, vol. 5, no. 2, pp. 615–620, Apr. 2024, doi: 10.52436/1.jutif.2024.5.2.1922.

[16] R. A. Ilma, S. Hadi, and A. Helen, "Twitter's hate speech multi-label classification using bidirectional long short-term memory (BiLSTM) method," in *2021 International Conference on Artificial Intelligence and Big Data Analytics*, IEEE, Oct. 2021, pp. 93–99, doi: 10.1109/ICAIBDA53487.2021.9689767.

[17] I. N. Yulita *et al*., "Bidirectional long short-term memory for analysis of public opinion sentiment on government policy during the COVID-19 pandemic," *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 11, 2023, doi: 10.14569/IJACSA.2023.0141189.

[18] D. R. Alghifari, M. Edi, and L. Firmansyah, "Implementasi bidirectional LSTM untuk analisis sentimen terhadap layanan Grab Indonesia," *Jurnal Manajemen Informatika (JAMIKA)*, vol. 12, no. 2, pp. 89–99, Sep. 2022, doi: 10.34010/jamika.v12i2.7764.

[19] A. Saputra, R. C. S. Hariyono, and N. M. Saraswati, "Analisis sentimen pengguna aplikasi MyPertamina menggunakan algoritma bidirectional long short-term memory," *Jurnal Eksplora Informatika*, vol. 13, no. 2, pp. 156–163, Mar. 2024, doi: 10.30864/eksplora.v13i2.973.

[20] S. Gomathi, K. N. Ram, and N. A. B. Mary, "Triplet encoded sequence-based membrane protein classification using BiLSTM," *Multimedia Tools and Applications*, vol. 83, no. 36, pp. 84251–84273, Apr. 2024, doi: 10.1007/s11042-024-19010-4.

[21] E. Khatizah and H.-S. Park, "Country-based COVID-19 DNA sequence classification in relation with international travel policy," *Applied Sciences*, vol. 14, no. 5, Art. no. 1916, Feb. 2024, doi: 10.3390/app14051916.

[22] E. Tabane, E. Mnkandla, and Z. Wang, "Harnessing ensemble deep learning for DNA sequence classification: Evaluating CNN, BiLSTM, and GRU architectures," *Preprints*, Jul. 2025, doi: 10.20944/preprints202507.0032.v1.

[23] R. Amelia, "Sentiment analysis of government policy in relocating the Republic of Indonesia's capital city," *Journal of Public Administration and Government*, vol. 5, no. 2, Aug. 2023. [Online]. Available: https://jurnal.fisip.untad.ac.id/index.php/JPAG

[24] A. Nugraha and M. U. Siregar, "Application of the Naive Bayes classifier method in the sentiment analysis of Twitter users about the capital city relocation," 2021.

[25] J. Muliawan and E. Dazki, "Sentiment analysis of Indonesia's capital city relocation using three algorithms: Naïve Bayes, KNN, and random forest," *Jurnal Teknik Informatika (JUTIF)*, vol. 4, no. 5, pp. 1227–1236, Oct. 2023, doi: 10.52436/1.jutif.2023.4.5.1436.

[26] Y. R. Dewi, N. W. S. Saraswati, M. O. E. Monny, I. B. G. Sarasvananda, and I. G. Andika, "Sentiment analysis of the relocation of the national capital on social media X," *Sinkron*, vol. 9, no. 2, pp. 625–636, Apr. 2025, doi: 10.33395/sinkron.v9i2.14622.

[27] A. W. Andi, C. Slamet, D. S. Maylawati, J. Jumadi, A. R. Atmadja, and M. A. Ramdhani, "Sentiment analysis of state capital relocation of Indonesia using convolutional neural network," in *2022 IEEE 8th International Conference on Computing, Engineering and Design (ICCED)*, IEEE, Jul. 2022, pp. 1–6, doi: 10.1109/ICCED56140.2022.10010503.

[28] M. D. Angelo, R. W. Harwenda, I. Budi, A. B. Santoso, and P. K. Putra, "Sentiment analysis and topic modeling of public opinion on Indonesia new capital city development policies," *Journal Eduvest*, vol. 5, no. 5, 2025. [Online]. Available: http://eduvest.greenvest.co.id

[29] S. Casola, I. Lauriola, and A. Lavelli, "Pre-trained transformers: An empirical comparison," *Machine Learning with Applications*, vol. 9, Art. no. 100334, Sep. 2022, doi: 10.1016/j.mlwa.2022.100334.

[30] V. Kumar, “Spatiotemporal sentiment variation analysis of geotagged COVID-19 tweets from India using a hybrid deep learning model,” *Scientific Reports*, vol. 12, no. 1, Art. no. 1849, Feb. 2022, doi: 10.1038/s41598-022-05974-6.
[31] A. Yenter and A. Verma, “Deep CNN-LSTM with combined kernels from multiple branches for IMDb review sentiment analysis,” in *2017 IEEE 8th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON)*, IEEE, Oct. 2017, pp. 540–546, doi: 10.1109/UEMCON.2017.8249013.
[32] X. Wang, W. Jiang, and Z. Luo, “Combination of convolutional and recurrent neural network for sentiment analysis of short texts,” 2016. [Online]. Available: https://github.com/ultimate010/crnn
[33] J. Wang, L.-C. Yu, K. R. Lai, and X. Zhang, “Dimensional sentiment analysis using a regional CNN-LSTM model,” in *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, Stroudsburg, PA, USA: Association for Computational Linguistics, 2016, pp. 225–230, doi: 10.18653/v1/P16-2037.
[34] G. Wiedemann, E. Ruppert, R. Jindal, and C. Biemann, “Transfer learning from LDA to BiLSTM-CNN for offensive language detection in Twitter,” 2018.
[35] S. Mallick *et al.*, “Deep hybrid CNN-LSTM framework for advanced social media sentiment analysis in data-driven marketing analytic,” *International Journal of Environmental Sciences*, pp. 169–180, Aug. 2025, doi: 10.64252/82p5f018.
[36] M. Hutagalung and Nurhasanah, “Media framing in the digital age: A narrative review of sociopolitical and cultural impacts,” *Communica: Journal of Communication*, vol. 3, no. 3, pp. 164–177, Jul. 2025.
[37] E. Musk, “Temporary limits on data scraping and post reading on X/Twitter,” Jan. 2025. [Online]. Available: https://x.com/elonmusk/status/1675187969420828672

Wanda Nugraha. Completed a bachelor's degree in mathematics at IPB University in 2025. Currently actively pursuing research and expertise in data science with research interests in machine learning, data analysis, and artificial intelligence.

Mochamad Tito Julianto. Lecturer in the Mathematics Study Program, SSMI, IPB University. Completed a bachelor's degree in mathematics at IPB and a master's degree in computer science at the University of Indonesia. Also serves as Chief Editor of MILANG Journal of Mathematics and Its Applications, contributing to the development of scientific publications in the field of mathematics and its applications. Research interests include mathematical computation, numerical modelling, and the application of mathematics to various real-world problems.

Mohamad Khoirun Najib. Lecturer in the Mathematics Study Program, SSMI, IPB University. Completed a bachelor's degree in mathematics and a master's degree in applied mathematics at IPB. Has extensive research publications in various national and international scientific journals, covering topics such as machine learning, data classification, time series models, and the application of computational methods in data analysis and real-world problems. Research interests include scientific computing, machine learning, numerical modelling, and applied data analysis.

Elis Khatizah. Completed bachelor's and master's degrees in Mathematics from IPB and ITB, as well as a doctorate in Computer Science from Ewha Womans University. Lecturer in the Mathematics Study Program, SSMI, IPB University, with a research focus on the application of machine learning and deep learning in biology and bioinformatics, particularly computational approaches to DNA and biological data analysis.