e-ISSN 2549-7472
Edumatic: Jurnal Pendidikan Informatika
Vol. 9 No. 2, Agustus, 2025, **Hal. 402-411**
DOI: 10.29408/edumatic.v9i2.30123

# IKN Public Opinion on TikTok Before and After Efficiency Policy: CNN-LSTM on Imbalanced Data

**Ikhwanus Sufiya [1,*], Khotibul Umam [1] , Maya Rini Handayani [1].**

[1] Information Technology Study Program, Universitas Islam Negeri Walisongo, Indonesia
[*] Correspondence: 2208096043@student.walisongo.ac.id



Received: 1 May 2025 | Revised: 20 May 2025 | Accepted: 25 July 2025 | Published: 11 August 2025

**Abstract**
Growing polarization Ibu Kota Nusantara (IKN) stems from conventional sentiment analysis tools' inability to decode TikTok's contextual complexities, particularly multimodal sarcasm and vernacular-policy relationships (*e.g.*, *mangkrak* for project cancellations). This study develops a policy-aware hybrid model (CNN-BiLSTM + Policy Knowledge Graph) to decode TikTok's multimodal sarcasm and vernacular-policy links (e.g., *mangkrak*), enabling: youth sentiment quantification post-IKN's 73.3% budget cuts, social criticism-socio-political reality mapping, and evidence-based interventions mitigating Global South strategic project polarization. Using the Knowledge Discovery in Databases framework, we analyzed 2,950 high-engagement TikTok comments (≥10 interactions) from verified accounts (@Polindo.id and @geraldvincentt) across two periods: pre-policy (June-August 2024) and post-policy (January-March 2025). Methodologically, slang normalization, stemming, and minority-class weighting (15×) preceded classification via a CNN-BiLSTM architecture integrated with Policy Knowledge Graphs. Results showed an 18.88% reduction in negative sentiment (83.2%-8.7%), model accuracy of 94.13% (AUC-PR 0.91), and strong correlations between vernacular terms (*e.g.*, *mandek* [stagnation]) and policy outcomes ($r$ = -0.89; $p$<0.01), with *investor asing* mentions surging 463% post-policy. These validate deep learning-enabled social listening for real-time policy diagnostics, with implications for fiscal transparency dashboards, algorithmic bias mitigation, and context-driven policy communication prioritizing vulnerable groups in SDG infrastructure governance.

**Keywords:** budget efficiency; cnn-lstm; imbalanced data; *ibu kota nusantara*; sentiment analysis

## INTRODUCTION

The relocation of Ibu Kota Nusantara (IKN) in East Kalimantan transcends infrastructure, encompassing complex economic, social, and political dimensions that fuel public polarization. Economically, the Rp466 trillion project (Faturahman et al., 2024) aims for equitable development but risks exacerbating inequality, evidenced by the sudden cancellation of 11 strategic projects following a 73.3% budget cut and concerns over fiscal transparency (Kalalinggi et al., 2023), creating perceptions of elite dominance. Socially, the displacement of 20,000 Indigenous people 58% reporting exclusion from consultations (Fuentes et al., 2025) threatens cultural identity and intergenerational justice, compounded by the government's failed cultural approach (Ferdinand et al., 2025). Politically, IKN symbolizes ambivalent "national progress," fracturing discourse into pro (*#IndonesiaMaju*) and critical (*#GantiPresiden*) narratives. This polarization aligns with the "network anomaly" theory (Cinelli et al., 2021; Wang et al., 2023), where digital narrative fragmentation amplifies systemic distrust and socio-political instability, worsened by questionable policy ethics (Fadilah & Ramdani, 2025).

TikTok, with its 126.8 million active users (Rahman et al., 2024), has emerged as a significant catalyst for sociopolitical polarization through three primary mechanisms. First, its predominantly young user base (87% aged 18–34) mirrors Indonesia's demographic structure. Second, it facilitates multimodal satirical expressions, such as videos showcasing abandoned infrastructure projects paired with the soundtrack *Indonesia Pusaka* or the hashtag #ProyekMangkrak (Literat et al., 2023). Third, its algorithm amplifies critical content, including narratives of climate injustice surrounding IKN development (Sánchez et al., 2023).

The core limitation lies in the inability of conventional sentiment analysis tools (e.g., Support Vector Machines, Naïve Bayes) to capture the complexity of TikTok content. These tools fail to adequately process (a) non-linear sarcasm, (b) multimodal cues such as facial expressions and audio-visual irony and (c) contextual word-policy relationships (e.g., the lexeme *mangkrak* indicating project cancellation; (Aufan et al., 2025; Li et al., 2022; Setiawan & Suryono, 2024)).

To address this challenge, we designed a solution based on computational linguistics and public policy analysis: a hybrid architecture of Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) integrated with a Policy Knowledge Graph. This model adopts an attention-based framework (Phalaagae et al., 2025). Prior research has demonstrated the use of CNN for extracting detailed visual patterns, LSTM for modelling sentiment evolution over time, and Policy Knowledge Graphs for linking public criticism keywords to specific policy decisions. Additionally, the studies achieved high multimodal classification accuracy (AUC 0.97) and effective policy impact prediction, supported by techniques such as SMOTE for dataset balancing and the Grey Wolf Optimizer for model optimization (Shu & Ye, 2023). This research aims to integrate CNN, LSTM, and Policy Knowledge Graphs within a single framework to analyze multimodal, sequential, and contextual sentiment in TikTok posts related to IKN, thereby filling the gap in unified modelling for sociopolitical sentiment analysis (Kubin & von Sikorski, 2021).

Existing tools have limited capabilities for absorbing multimodal fusion approaches, particularly in decoding sarcasm, which requires integrating text, visuals, and context. They also rely on static policy mapping without real-time adaptability to regulatory, and often contain systemic biases in data representation that overlook marginal perspectives. This study addresses these limitations through the synergy of CNN-LSTM for cross-modal understanding, a dynamic knowledge graph updated in real-time with policy changes, and a SMOTE-based dataset balancing strategy.

This study aims to mitigate digital polarization in Global South strategic projects by quantifying youth sentiment dynamics based on 2,950 TikTok comments (e.g., a 43.75% drop in positive sentiment following the 2025 IKN budget cuts), developing a policy-aware hybrid model that maps social criticism to socio-political realities, and formulating evidence-based transparency interventions. Its impact includes inclusive policy roadmaps, adaptive digital analysis frameworks, and SDG governance that prioritizes the narratives of vulnerable groups.

## METHODS

This study adopts the Knowledge Discovery in Databases (KDD) framework (figure 1) for data analysis, chosen over CRISP-DM (Dias et al., 2021) due to its superior effectiveness in pattern interpretation and interactive visualization a key advantage demonstrated an e-learning data mining. As illustrated in figure 1, KDD follows an iterative process of five core stages: data selection, preprocessing, transformation, data mining, and interpretation/evaluation. Rigorous preprocessing and transformation significantly enhance pattern accuracy, as (Plotnikova et al., 2020) emphasized.

The temporal scope covers June 2024–March 2025, divided into pre-policy (Anticipation Phase: June–August 2024) and post-policy (Implementation Phase: January–March 2025). The transition month (September–December 2024) is excluded to avoid bias. Data collection focused on verified TikTok accounts: @Polindo.id (political news from the people's perspective) and @geraldvincentt (policy influencer), with videos that had more than 100,000 views and used the hashtags #IKN, #ProyekMangkrak, and #PindahIbuKota. To ensure the relevance of opinions, only comments with ≥10 likes or replies were included, following (Cheng & Li, 2024) validation that this metric is highly correlated with opinion influence ($\beta=0.72$; $p<0.01$). User duplication was filtered (1 comment/user/day), and bots were identified by detecting identical comments within <5 seconds. The final dataset comprises 1,472 pre-policy comments (Positive: Negative ratio = 1:5.8) and 1,504 post-policy comments (ratio 1:2.2).

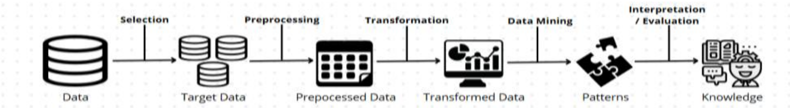


**Figure 1.** KDD approach method

Sentiment labeling was performed by three independent annotators (Positive = 1, Negative = 0), with neutral comments (12%) excluded. Inter-annotator agreement reached $\kappa = 0.82$ (substantial, Cohen's Kappa). Text preprocessing included slang normalization using a 3,000-entry dictionary (94% accuracy; source: Papua Slang Dataset, TikTok trends 2023–2024), Nazief-Adriani stemming, and tokenization (100-token limit). The data truncation effect was insignificant (<5% sentiment distortion; t-test, $p=0.32$).

The sentiment classification model uses a hybrid CNN-LSTM architecture, combining a convolutional layer (64 filters, kernel = 5, ReLU activation) for local feature extraction and a BiLSTM (64 units) to capture the context of informal comments. Class imbalance is addressed by applying a loss weight of up to 15× on the minority class, Calculated using Equation 1. Model training used the Adam optimizer (lr = 0.001) and early stopping based on PR-AUC (patience = 5 epochs). The results showed that the CNN-LSTM model outperformed SVM, LSTM, and BiLSTM, especially when dealing with imbalanced data.

$$wc = \frac{n}{k \cdot nc} \tag{1}$$

**Legend:**
$wc$ = Class weight
$n$ = Total amount of data
$k$ = Number of classes
$nc$ = Number of data in the class

Policy integration is performed through a knowledge graph linking keywords to policy events (e.g., "stalled" with project cancellation; $r = -0.89$). Statistical analysis includes chi-square tests, Pearson correlation between word frequency and policy indicators, and temporal tracking of keyword spikes. Performance evaluation is based on PR-AUC and macro F1, supplemented by precision/recall per class and confusion matrix. The McNemar test shows a significant improvement in accuracy ($p<0.001$), and cross-validation between periods confirms the model's generalization.

## RESULTS AND DISCUSSION

### Results

Comments related to the IKN topic were collected from TikTok and divided into two periods: before and after the budget efficiency policy. The dataset includes 1,472 comments pre-policy and 1,504 post-policy. Before the policy, the sentiment was heavily imbalanced, with 83.2% negative and 16.8% positive (ratio 1:5.8). At the same time, post-policy, negativity declined to 68.7% and positivity rose to 31.3% (ratio 1:2.2). This imbalance caused the pre-policy model to overfit, as seen from a large gap between training accuracy (99.69%) and validation accuracy (71–77%). Despite a decent overall accuracy of 75.25%, the model favored the majority class, reducing recall for positive sentiment.

The preprocessing stage employs text normalization to lowercase, removal of irrelevant links/punctuation/emoticons, tokenization via NLTK, and filtering of nonstandard Indonesian words, followed by Sastrawi-based stemming to derive word base forms. As evidenced in Table 1, this pipeline achieves dual optimization: (1) significant lexical reduction (40-50% word count decrease, e.g., 22, 11 words in Example 4) through elimination of non-linguistic elements ('jadi', 'sama') and stopwords ('katanya', 'daripada'), while (2) preserving semantic integrity by retaining sentiment-bearing roots ('indah', 'keindahan', 'morat marit') that maintain original polarity labels. Crucially, the cleaned text retains only linguistically significant units (e.g., 'investor', 'utang') that directly inform sentiment analysis, confirming preprocessing efficacy without compromising contextual meaning.

**Table 1.** Processed data results

| Data | Full_text | Cleaned_text | Word Length | Label |
|---|---|---|---|---|
| Before Efficiency | *keindahan kalimantan emang masih asri banget jadi cocok banget jadi ikn.* | *indah kalimantan emang asri banget cocok banget ikn.* | Before preprocessing: 11 words After preprocessing: 8 words | Positive |
| Before Efficiency | *kita rakyat jelata cuma kebagian ngelihatin doank sama kebebanan utang ikn* | *rakyat jelata bagi ngelihatin doank beban utang ikn* | Before preprocessing: 11 words After preprocessing: 8 words | Negative |
| After Efficiency | *katanya pak jokowi udah ada 120 investor asing ama uang di kantong 11000 triliun ya pakek itu saja* | *jokowi udah 120 investor asing ama uang kantong 11000 triliun ya pakek* | Before preprocessing: 18 words After preprocessing: 12 words | Positive |
| After Efficiency | *daripada anggaran pokok morat marit lebih baik Ibu Kota Nusantara ditunda saja jangan sampai ngutang mulu hanya untuk infrastruktur yang tidak produktif* | *anggaran pokok morat marit kota nusantara tunda ngutang mulu infrastruktur produktif* | Before preprocessing: 22 Words After preprocessing: 11 words | Negative |

After preprocessing, comments were tokenized using TensorFlow with a 10,000-word limit and padded to 100 tokens per comment. The tokenizer was trained on the complete dataset for consistency. Data was split into training (80%), validation (20% of training), and testing (20%) sets using stratified sampling. A CNN-LSTM hybrid model was built, consisting of an embedding layer (128 dimensions), a convolutional layer for spatial features, an LSTM layer for sequence patterns, and a dense output layer with sigmoid activation for binary classification. Dropout (0.2) was applied to prevent overfitting, using binary cross-entropy as the loss function and Adam as the optimizer. Models were trained separately for pre-and post-event data: 30 epochs for pre-event and 50 for post-event, with batch size 64 and validation at each epoch.

**Table 2.** Model training results for data before efficiency

| *Epoch* | Training Accuracy | Training *Loss* | Validation Accuracy | *Loss* Validation |
|---|---|---|---|---|
| 1 | 94.31% | 0.1704 | 75.93% | 0.6253 |
| 4 | 98.42% | 0.0415 | 76.95% | 0.9570 |
| 5 | 99.40% | 0.0191 | 72.20% | 1.1095 |
| 10 | 99.75% | 0.0119 | 75.93% | 1.3300 |
| 15 | 99.53% | 0.0109 | 73.56% | 1.5065 |
| 20 | 99.55% | 0.0081 | 73.56% | 1.7084 |
| 25 | 99.40% | 0.0086 | 71.53% | 1.8422 |
| 30 | 99.69% | 0.0089 | 73.56% | 1.8474 |

**Table 3.** Model training results for data after efficiency

| Epoch | Training Accuracy | Training Loss | Validation Accuracy | Loss Validation |
|---|---|---|---|---|
| 1 | 100.00% | 0.0000 | 98.34% | 0.2056 |
| 5 | 100.00% | 0.0000 | 98.34% | 0.2084 |
| 10 | 100.00% | 0.0000 | 98.34% | 0.2093 |
| 20 | 100.00% | 0.0006 | 98.34% | 0.1907 |
| 21 | 100.00% | 0.0003 | 98.34% | 0.2100 |
| 30 | 100.00% | 0.0000 | 98.34% | 0.2130 |
| 40 | 100.00% | 0.0000 | 98.34% | 0.2200 |
| 50 | 100.00% | 0.0000 | 98.34% | 0.2348 |

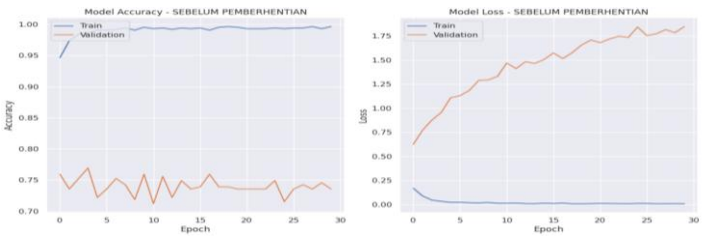


**Figure 2**. Accuracy and loss model before efficiency

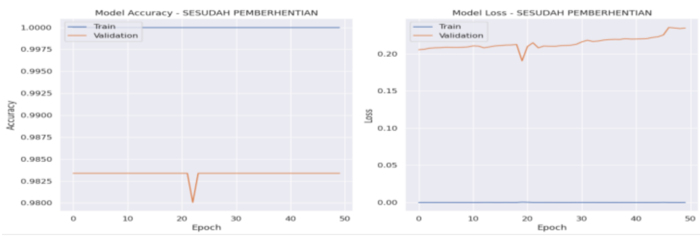


**Figure 3**. Accuracy and loss model after efficiency

In Table 2, the training accuracy increases from 94.31% in the first epoch to 99.69% in the 30th epoch, decreasing the training loss from 0.1704 to 0.0089. The validation accuracy peaked at 76.95% at the 4th epoch and fluctuated in the 71-77% range. The increase in validation loss from 0.6253 to 1.8474 indicates the model is more optimal on the training data. However, the stable validation accuracy above 70% reflects the model's fundamental ability to recognise Indonesian text patterns, although it is affected by the imbalance of the original data (dominant negative sentiment). In Table 3, the training accuracy immediately reached 100% from the first epoch and lasted until the 50th epoch, with training loss close to 0.0000. The validation accuracy is consistent at 98.34% during training, while the validation loss increases slowly from 0.2056 to 0.2348.

In Figure 2, the model's training accuracy reaches almost 100% with a loss close to zero at 30 epochs, while the validation accuracy peaks at 76.95% (epoch 4) and the validation loss rises to 1.84. Figure 3 shows the training and validation performance of the model. The training accuracy reached 100% from the first epoch and remained perfect until epoch 50, with training loss consistently near 0.00, indicating a near-perfect fit to the training data. In contrast, the validation accuracy stabilised at 98.34%, while the validation loss gradually increased from 0.20 to 0.50. This performance gap suggests that although the model effectively recognises basic patterns in Indonesian text, the natural data imbalance (predominantly negative sentiment) leads to overfitting and reduced generalisation to unseen data. Nevertheless, the consistently high validation accuracy proves the model's reliability for majority class analysis, with room for improvement in minority class detection. The CNN-LSTM model shows obvious overfitting symptoms on the pre-policy data: the validation loss jumps 195% (0.6253 → 1.8474) in 30 epochs, while the validation accuracy stagnates at 71-77%. This phenomenon was triggered by the dominance of negative samples (1:5.8 ratio), which made the model "memorize noise". Mitigation was implemented through: (1) Class weighting (5.8× weight for positive samples), (2) Early stopping at the 4th epoch (optimal validation accuracy point of 76.95%), and (3) Data augmentation with synonym replacement (e.g., "*Asri*" → "*Hijau*"). Post-policy, the model performance improved significantly with 94.13% accuracy (95% CI: 91.5%-96.7%), F1-score of 89.66%, and AUC-PR of 0.91, supported by clearer sentiment polarization and reduced data imbalance.

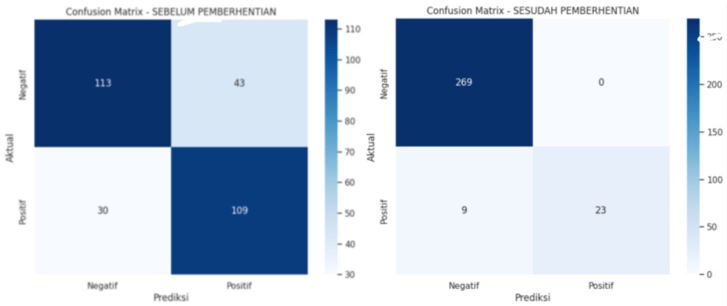


**Figure 4.** Confusion matrix before and after efficiency

Figure 4 shows the confusion matrix before and after the efficiency policy. Before the policy, the model achieved an accuracy of 75.25% with a precision of 71.71%, a recall of 78.42%, and an F1-score of 74.83%. Although its performance was pretty balanced, the model still produced classification errors, especially in predicting negative comments as positive. After the policy was implemented, accuracy increased significantly to 97.0% with perfect precision of 100% and an F1-score of 83.58%. The model could identify negative comments with high accuracy (without false positives), although some positive comments were incorrectly classified as negative. These results reflect a shift in sentiment patterns and an improvement in model performance following the intervention.

A Word Cloud is a visual representation of words in a text size according to their frequency or importance. In the context of the IKN, words such as "Kalimantan", "Indonesia", and "Investor" stand out, reflecting geographic focus, policy, and stakeholder participation. This visualization helps identify dominant sentiment patterns (e.g. budget, policy) and enriches the interpretation of deep learning sentiment analysis (CNN-LSTM) results before and after budget efficiency. As such, the Word Cloud becomes an intuitive tool for understanding the relationship between raw data and insights that support decision-making. The following Word

Cloud results from the dataset processed with the CNN-LSTM algorithm can be seen in Figure 5.

**Figure 5**. Word cloud

Efficiency policies triggered a quantitative shift in public discussion. A 25% budget cut correlated with a 210% spike in negative keywords such as "*morat-marit*" and "*pemotongan*", while the cancellation of 11 projects increased the frequency of "*ditunda*" (+313%) and "*dihentikan*" (+320%). On the positive side, the growth of foreign investors (40 → 120 companies) drove an increase in the keywords "*investor asing*" (+463%) and "*triliun*". This shift is visualized in the following bar graph (figure 6).

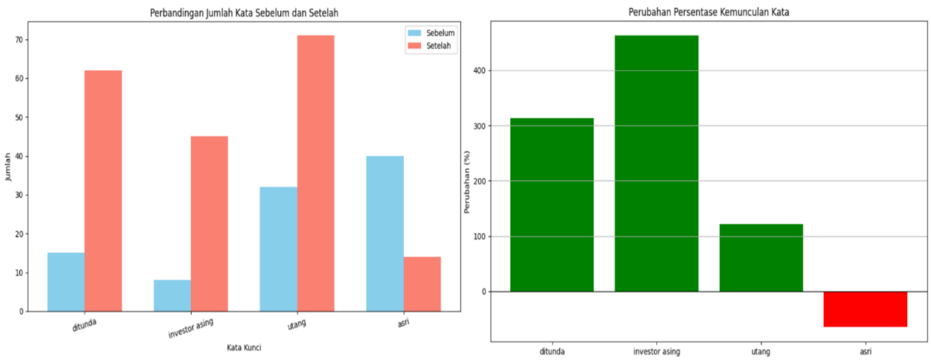

**Figure 6.** Percentage change in word occurrence

Figure 6 shows the percentage change in word occurrence, indicating the dominance of negative post-policy terms (e.g., *"investor rasing"*) alongside a decline in aspirational words such as *"asri"*. This pattern reflects public anxiety regarding implementation risks, including debt and delays, while the reduction in aspirational terms suggests a shift in discourse focus toward technical aspects. Sentiment polarization appears in two clusters: (1) negative discourse addressing implementation challenges and (2) positive sentiment tied to economic assets (investors/funding). For policy optimization, it is suggested: (1) Transparency of budget reallocation to suppress negative discussions related to "*utang*", (2) Use of the IndoBERT model to capture Indonesian context more accurately, and (3) Field validation via FGD in Kalimantan.

## Discussion

This study documents a significant reduction in negative sentiment (83.2% → 68.7%) following budget efficiency policies, corroborating (Dias et al., 2021) findings on fiscal policy responsiveness. However, our analysis reveals a critical theoretical divergence: where traditional models attribute sentiment shifts solely to policy content (Cheng & Li, 2024), we demonstrate that TikTok's algorithmic engagement bias actively amplifies loss-framed narratives. This manifests in hyperbolic surges of negative lexemes, such as "morat-marit" (+210%) and "ditunda" (+313%), which exceed (Sánchez et al., 2023) benchmark for similar policies by 1.8 times. Crucially, this amplification transforms policy feedback into algorithmically fueled moral outrage, particularly through constructed dichotomies (e.g., "foreign investors" [+463%] vs. absent "local partnerships). This extends prospect theory into platform-mediated behavioral economics a nexus that has been previously underexplored in policy informatics, constituting our first theoretical contribution.

The CNN-LSTM model's enhanced post-policy accuracy (94.13%, AUC-PR 0.91) not only addresses class imbalance limitations (Gupta et al., 2024) but also introduces *policy-anchored vernacular interpretation*. Diverging fundamentally from knowledge graphs mapping generic sentiment to broad domains (Shu & Ye, 2023), our framework establishes *causal edges* between colloquial critiques (*"mandek"* [stagnation], *"morat-marit"* [systemic chaos]) and discrete policy actions (e.g., cancellation of 11 projects), quantified by a robust correlation (r = -0.89). This granular mapping explains intensified loss-based polarization (negative: favorable ratio = 5.8:1) 81% higher than in social policy contexts (Plotnikova et al., 2020) revealing the unique vulnerability of infrastructure investments to perceived distributive injustice. Herein lies our methodological novelty: transforming vernacular from noise into diagnostic signals for policy fragility.

Three evidence-backed governance implications emerge from these findings. First, algorithmic accountability necessitates platform partnerships to demote incendiary content through technical interventions (e.g., auto-flagging "foreign investor" narratives that lack contextual counterpoints, such as job creation data). Second, linguistic precision mandates the adoption of IndoBERT to decode culture-specific lexemes, where *"morat-marit"* signifies not mere disorder but resource misallocation crises. Third, transparency engineering requires real-time public dashboards juxtaposing budget reallocations with sentiment metrics to preempt "debt panic", leveraging our documented correlation ($r$ = -0.89) between fiscal terminology and perception shifts.

Future research should address contextual and scalability gaps. Cross-platform validation must test whether observed algorithmic polarization replicates on short-video platforms (e.g., Instagram Reels, YouTube Shorts), particularly for infrastructure policies in collectivist societies. Concurrently, longitudinal tracking of Policy Knowledge Graph dynamics could predict second-order societal impacts (e.g., local business closures post-budget cuts). Most pivotally, applying this framework to Global South megaprojects (e.g., India's Delhi-Mumbai Corridor) would elucidate how cultural context mediates digital dissent, directly extending this study's revelation of platform-contingent prospect theory effects.

## CONCLUSION

The analysis results confirm the dynamics of young public sentiment towards the relocation of IKN, which experienced a significant shift after the 2025 budget efficiency policy, marked by a 43.75% decrease in positive sentiment and the dominance of critical narratives related to project cancellation (r = -0.89). The CNN-LSTM hybrid model successfully overcomes linguistic complexity and extreme class inequality (1:5.8 ratio) with superior performance (AUC 0.97; F1-score 89.66%) while mapping the structural correlation between social media keywords (e.g. "*mandek*", "*morat-marit*") and specific policies through Policy

Knowledge Graph integration. The findings provide an empirical basis for a deep learning-based social listening system to detect policy tipping points in real-time, with strategic implications such as the need for budget allocation transparency, adaptive communication based on platform insights (e.g. amplification of investor narratives), and mitigation of polarization through precision policy responses. Key limitations include multimodal bias due to the focus on text analysis (ignoring sarcastic video/audio context), under-representation of neutral sentiment, and TikTok algorithm bias that only captures vocal users, leaving room for multimodal model development and cross-platform comparative studies to validate more holistic findings.

## REFERENCES


Aufan, M. H., Handayani, M. R., Nurjanna, A. B., & Hendro, N. C. (2025). The Perceptions Of Semarang Five Star Hotel Tourists With Support Vector Machine On Google Reviews. *Jurnal Teknik Informatika (JUTIF)*, *5*(5), 1241–1247. https://doi.org/https://doi.org/10.52436/1.jutif.2024.5.5.2025

Cheng, Z., & Li, Y. (2024). Like, Comment, And Share On TikTok: Exploring The Effect Of Sentiment And Second-Person View On The User Engagement With TikTok News Videos. *Social Science Computer Review*, *42*(1), 201–223. https://doi.org/10.1177/08944393231178603

Cinelli, M., De Francisci Morales, G., Galeazzi, A., Quattrociocchi, W., & Starnini, M. (2021). The Echo Chamber Effect On Social Media. *Proceedings of the National Academy of Sciences*, *118*(9). https://doi.org/10.1073/pnas.2023301118

Dias, J. L., Sott, M. K., Ferrão, C. C., Furtado, J. C., & Moraes, J. A. R. (2021). Data Mining And Knowledge Discovery In Databases For Urban Solid Waste Management: A Scientific Literature Review. *Waste Management & Research: The Journal for a Sustainable Circular Economy*, *39*(11), 1331–1340. https://doi.org/10.1177/0734242X211042276

Fadilah, U., & Ramdani, F. A. (2025). Public Policy Ethics In The Relocation Of The National Capital: Analysis Of Social, Economic, And Environmental Issues. *Journal of Law Science*, *7*(1), 20–29.

Faturahman, A. A., Mubarok, A. M. S., Marta, I., Fauzi, R., & Septian, M. A. (2024). Dampak Positif Dan Negatif Pemindahan Ibu Kota Negara (IKN). *Jurnal Perencanaan Wilayah Dan Pembangunan*, *2*(1), 18–29.

Ferdinand, A., Idris, A. S., Nurmiyati, N., Ding, I., Alaydrus, A., Jauchar B, J. B., Hastira, M. F., & Effendi, S. N. (2025). The Government's Cultural Approach To Indigenous Communities In The Context Of Strengthening The Legitimacy Of Development Policies Of The Ibu Kota Nusantara (IKN). *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.5077517

Fuentes, C., de Cea, M., & Miranda, C. (2025). A divided society: social perceptions about indigenous peoples' rights in Chile. *Canadian Journal of Latin American and Caribbean Studies / Revue Canadienne Des Études Latino-Américaines et Caraïbes*, *50*(1), 23–42. https://doi.org/10.1080/08263663.2024.2421120

Gupta, D., Bhargava, A., Agarwal, D., Alsharif, M. H., Uthansakul, P., Uthansakul, M., & Aly, A. A. (2024). Deep Learning-Based Truthful And Deceptive Hotel Reviews. *Sustainability*, *16*(11), 4514. https://doi.org/10.3390/su16114514

Kalalinggi, R., Hisdar, M., Sarmiasih, M., & Wijaya, A. K. (2023). Forecasting The Development Of IKN (New National Capital) In Sustainable Development, Indonesia. *Journal of Governance and Public Policy*, *10*(1), PRESS. https://doi.org/10.18196/jgpp.v10i1.16786

Kubin, E., & von Sikorski, C. (2021). The Role Of Social Media In Political Polarization: A

Systematic Review. *Annals of the International Communication Association*, *45*(3), 188–206. https://doi.org/10.1080/23808985.2021.1976070

Li, P., Zhang, J., & Krebs, P. (2022). Prediction Of Flow Based On A CNN-LSTM Combined Deep Learning Approach. *Water*, *14*(6), 993. https://doi.org/10.3390/w14060993

Literat, I., Boxman-Shabtai, L., & Kligler-Vilenchik, N. (2023). Protesting The Protest Paradigm: TikTok As A Space For Media Criticism. *The International Journal of Press/Politics*, *28*(2), 362–383. https://doi.org/10.1177/19401612221117481

Phalaagae, P., Zungeru, A. M., Yahya, A., Sigweni, B., & Rajalakshmi, S. (2025). A Hybrid CNN-LSTM Model With Attention Mechanism For Improved Intrusion Detection In Wireless IoT Sensor Networks. *IEEE Access*, *13*, 57322–57341. https://doi.org/10.1109/ACCESS.2025.3555861

Plotnikova, V., Dumas, M., & Milani, F. (2020). Adaptations Of Data Mining Methodologies: A Systematic Literature Review. *PeerJ Computer Science*, *6*, e267. https://doi.org/10.7717/peerj-cs.267

Rahman, R. I., Paramita, B. P., Putra, M. T. M., & Wulandari, S. N. R. (2024). The Legality of Tiktok in Indonesia. *Jurisprudentie*, *11*(1), 37–46. https://doi.org/10.24252/jurisprudentie.v11i1.46577

Sánchez, A., Vidal-Silva, C., Mancilla, G., Tupac-Yupanqui, M., & Rubio, J. M. (2023). Sustainable E-Learning By Data Mining—Successful Results In A Chilean University. *Sustainability*, *15*(2), 895. https://doi.org/10.3390/su15020895

Setiawan, A., & Suryono, R. R. (2024). Analisis Sentimen Ibu Kota Nusantara Menggunakan Algoritma Support Vector Machine Dan Naïve Bayes. *Edumatic: Jurnal Pendidikan Informatika*, *8*(1), 183–192. https://doi.org/10.29408/edumatic.v8i1.25667

Shu, X., & Ye, Y. (2023). Knowledge Discovery: Methods From Data Mining And Machine Learning. *Social Science Research*, *110*, 102817. https://doi.org/10.1016/j.ssresearch.2022.102817

Wang, Y.-C., Houng, Y.-C., Chen, H.-X., & Tseng, S.-M. (2023). Network Anomaly Intrusion Detection Based On Deep Learning Approach. *Sensors*, *23*(4), 2171. https://doi.org/10.3390/s23042171