Eduvest – Journal of Universal Studies
Volume X Number X, Mei, 2025
p- ISSN 2775-3735- e-ISSN 2775-3727

SENTIMENT ANALYSIS AND TOPIC MODELING OF PUBLIC OPINION ON INDONESIA NEW CAPITAL CITY DEVELOPMENT POLICIES

Michael David Angelo[1*], Reyhan Widyatna Harwenda[2], Indra Budi[3], Aris Budi Santoso[4], Prabu Kresna Putra[5]
Faculty of Computer Science, Universitas Indonesia[1,2,3,4]
National Research and Innovation Agency, Indonesia [5]
Email : michael.david31@ui.ac.id, reyhan.harwenda@ui.ac.id , indra@cs.ui.ac.id, aris.budi@ui.ac.id, prab003@brin.go.id

ABSTRACT

This study explores public sentiment and thematic concerns surrounding Indonesia's Ibu Kota Nusantara (IKN) development by analyzing 9,451 tweets from 2017 to 2025. The research uses sentiment analysis, topic modeling, and statistical testing to assess public opinion dynamics during two political phases: the administration of President Joko Widodo and the transition to Prabowo Subianto. The study compares traditional machine learning models (Naïve Bayes, SVM, AdaBoost, XGBoost, LightGBM) with the IndoBERT transformer model for sentiment classification. IndoBERT outperformed other models, showing higher accuracy, precision, and recall. The sentiment shift analysis reveals significant differences between the two political phases, highlighting the influence of leadership transitions on public perception. Topic modeling uncovers dominant themes such as environmental sustainability, socio-economic impacts, governance, and infrastructure, providing insights into public concerns. These findings emphasize the need for tailored communication strategies during political transitions and underscore the importance of responsive governance in large-scale policy initiatives. The research contributes to the field by demonstrating the effectiveness of advanced machine learning in understanding public opinion on major national projects and offers actionable insights for policymakers.

KEYWORDS *Sentiment Analysis, Topic Modeling, Public Opinion, IndoBERT, Machine Learning, Twitter, Political Transition, IKN.*

Article Info:

Submitted: 05-05-2025	Final Revised: 14-05-2025	Accepted: 19-05-2025	Published: 22-05-2025

How to cite:	Michael David Angelo, et al. (2025). Sentiment Analysis and Topic Modeling of Public Opinion on Indonesia New Capital City Development Policies. Journal Eduvest. *5*(5): 5951-5969
E-ISSN:	2775-3727

INTRODUCTION

The relocation of Indonesia's capital city from Jakarta to East Kalimantan designated as Ibu Kota Nusantara (IKN) marks a transformative national initiative aiming to promote sustainable development, reduce the burden on Jakarta, and catalyze economic growth in underdeveloped regions. However, the policy has sparked widespread debate and polarized public opinion, ranging from optimistic support to vocal criticism, particularly on environmental, socio-political, and financial grounds [1].

As the discourse surrounding IKN evolves, social media has become a vital lens for capturing real-time public sentiment and identifying shifts in perception. Platforms like Twitter and YouTube not only reflect grassroots reactions but also serve as feedback mechanisms that can guide inclusive policy development [2]. Sentiment analysis and topic modeling, two key Natural Language Processing (NLP) techniques, offer valuable insights into how citizens perceive and react to complex public policies over time [3]. Sentiment analysis can provide an overview of the current distribution of public sentiment, while topic modeling can uncover the key themes and issues being discussed. These insights are important for policymakers to understand the evolving dynamics of public opinion and respond more effectively to public concerns.

Previous studies have leveraged sentiment analysis in high-stakes policy contexts, such as infrastructure megaprojects [2], fuel subsidy reforms [4], and presidential elections [5]. These works underline the value of combining machine learning with social media analytics to identify opinion trends, validate policy acceptance, and support evidence-based governance. Moreover, the use of different machine learning models—from classical algorithms like Naïve Bayes and Support Vector Machine to advanced transformers like BERT—has enabled comparative evaluations of performance in sentiment classification [6].

In the case of IKN, the sentiment landscape is particularly dynamic. The transition of political leadership from President Joko Widodo to his successor Prabowo Subianto may influence how citizens perceive the continuity and legitimacy of IKN development. Detecting such sentiment shifts through statistical tests like the independent t-test can provide empirical insights into public response across political periods, a perspective that is still underexplored in current literature. Therefore, this study will examine IKN's public sentiment and find the most discussed topic in two distinct political phases: the era of President Joko Widodo and the subsequent phase under Prabowo Subianto. This approach allows for a more nuanced understanding of how changes in political leadership affect public opinion on a major national policy.

This study investigates public opinion on Indonesia's IKN development by analyzing social media data through sentiment analysis and topic modeling, employing a mixed-method approach that combines machine learning classification, statistical analysis, and topic extraction to contribute methodologically and practically by (1) identifying optimal sentiment analysis models for Indonesian contexts, (2) analyzing public opinion dynamics over time, and (3) uncovering key public concerns and expectations around the policy. Previous research highlights the effectiveness of NLP and ML techniques, such as sentiment analysis and topic modeling, in assessing public sentiment toward government policies, including infrastructure projects in Indonesia, by classifying emotions and extracting themes from unstructured social media data to gauge public reactions.

Sentiment analysis has been extensively employed to assess public sentiment regarding major governmental projects. For instance, Nokkaew et al. (2024) analyzed public sentiment on infrastructure projects such as the Thailand-China high-speed train and Laos-China railway projects, revealing sentiment shifts that occurred in different project phases, influenced by governmental communications and media portrayals [7]. Similarly, Alotaibi and Nadeem (2024) applied sentiment analysis to social media data to explore public opinion on educational reforms in Saudi Arabia, highlighting the utility of platforms like Twitter in capturing real-time sentiment [8]. These studies underscore how sentiment shifts can mirror the impacts of governmental messaging and the public's evolving attitudes toward such large-scale projects.

Topic modeling, in conjunction with sentiment analysis, is widely used to uncover the key themes driving public sentiment in complex policy domains. Al-Jundi et al. (2023) utilized a hybrid approach combining machine learning and lexicon-based techniques to analyze sentiment related to government policies, emphasizing the role of nuanced responses in issues such as AI and education [9]. In the case of Indonesia's IKN policies, this methodology is particularly relevant, as sentiment is driven by a wide range of socio-political, environmental, and economic concerns.

The rapid advancement of machine learning techniques has significantly enhanced the effectiveness of sentiment analysis. For example, Sharma et al. (2020) introduced an intelligent learning-based opinion-mining model for governmental decision-making, demonstrating the utility of both supervised and deep learning methods in classifying public sentiment on government policies [3]. Furthermore, studies such as those by Alotaibi and Nadeem (2024) have shown that fine-tuned transformer models like AraBERT can outperform traditional classifiers like Support Vector Machines (SVM) in sentiment classification, offering more nuanced insights into public opinion [10]. This aligns with the broader trend of utilizing transformer-based models like BERT and RoBERTa, which have shown superior performance in capturing public sentiment on social media [11], [12].

In the Indonesian context, where platforms such as Twitter are central to public discourse, traditional models such as SVM and Naïve Bayes have been effectively used for sentiment classification [4]. Zaki Al Faridzi et al. (2023) applied SVM to analyze public sentiment regarding Indonesia's fuel price hikes, achieving high classification accuracy with the use of techniques like SMOTE [4]. Such studies provide a foundation for selecting the most appropriate machine learning models for sentiment analysis in the Indonesian context, particularly for the IKN project.

Social media platforms, especially Twitter, have become essential tools for capturing public sentiment in real time. Jaiswal et al. (2023) examined ESG investing sentiments on Twitter, applying topic modeling to uncover themes like climate change risks and corporate governance issues. This study highlighted the potential of social media to offer valuable, real-time insights into public sentiment on a wide range of topics [13]. For Indonesia's IKN policies, Twitter serves as a critical space for both support and criticism, making sentiment analysis a powerful tool for policymakers to gauge public reaction and adjust strategies accordingly.

Comparative studies of sentiment analysis models are crucial in determining the most effective techniques for different applications. For instance, Zainulabdeen et al. (2024) compared various machine learning models for sentiment classification of COVID-19 vaccine opinions on Twitter, demonstrating that transformer-based models like BERT outperform traditional models such as Naïve Bayes in terms of accuracy and recall [6]. These findings suggest that while simpler models such as Naïve Bayes are effective for straightforward sentiment classification tasks, more advanced transformers like BERT and RoBERTa provide superior performance when dealing with complex datasets such as those related to the IKN project.

Recent advancements in ensemble-based machine learning methods, such as XGBoost, AdaBoost, and LightGBM, have significantly enhanced the capabilities of sentiment analysis, particularly in processing large-scale social media data. For instance, a study by Medvedeva and Al-LAMI (2023) demonstrated that integrating Convolutional Neural Networks (CNN) with XGBoost improved sentiment classification accuracy on Twitter data by 4–5% [14], outperforming traditional classifiers. Similarly, Wang (2025) applied XGBoost in conjunction with word vector representations to analyze mobile phone reviews, achieving higher accuracy and recall compared to models relying solely on large language models [15]. In the realm of social media analytics, ensemble approaches combining deep learning architectures with XGBoost have shown promising results; Ling et al. (2022) proposed a hybrid model integrating BiLSTM, Attention mechanisms, CNN, and XGBoost, which achieved over 95% accuracy in sentiment classification tasks across English and Chinese datasets [16].

Other ensemble methods such as AdaBoost have also demonstrated notable effectiveness in sentiment classification tasks. Hasan et al. (2021) found that AdaBoost outperformed simpler classifiers, especially in datasets characterized by class imbalance and linguistic variability, which are common in social media discourse. Furthermore, LightGBM, renowned for its computational efficiency and effectiveness in handling large-scale textual data, has been successfully employed in sentiment analysis studies. Onan (2021) highlighted LightGBM's capability to deliver robust sentiment classification performance, particularly in handling extensive social media datasets, surpassing traditional classifiers like Logistic Regression and Support Vector Machines [17].

Ensemble-based machine learning methods present viable alternatives or supplements to transformer models like BERT and RoBERTa, offering robust tools for capturing nuanced sentiment patterns in Indonesian public discourse, particularly regarding policies such as the IKN project, thereby enhancing opinion analysis and supporting data-driven policymaking. However, despite sentiment analysis's widespread use, few studies have examined the evolving nature of public sentiment toward Indonesian policies like IKN, leaving gaps in comparing simpler models with advanced transformers in local contexts and assessing how political transitions—such as from Joko Widodo to Prabowo Subianto—influence public sentiment. Addressing these gaps would yield critical insights into sentiment dynamics and improve policy responsiveness.

This research offers a novel approach to analyzing public sentiment on Indonesia's Ibu Kota Nusantara (IKN) development by incorporating sentiment analysis, topic modeling, and sentiment shift analysis across political transitions. Unlike previous studies focusing on stable political periods, it examines sentiment changes resulting from the political shift from President Joko Widodo to Prabowo Subianto, adding a comparative dimension to understanding how leadership transitions affect public opinion on national policies like IKN

METHOD

This study employs sentiment analysis, topic modeling, and statistical testing to analyze public opinion on Indonesia's IKN development. The methodology consists of key stages: data collection, labeling, preprocessing, sentiment classification, sentiment shift analysis, and topic modeling (as illustrated in Figure 2). Data was collected by scraping posts from X (formerly Twitter) using Python's Tweet Harvest library on Google Colab, with CSV exports for processing. The choice of X was supported by a review of 22 peer-reviewed studies on public sentiment toward government policies, 11 of which specifically used Twitter/X, highlighting its effectiveness in capturing public discourse. The dataset spans April

9, 2017, to April 25, 2025, enabling longitudinal analysis across two political phases: Phase 1 (Jokowi's Administration: April 9, 2017 – October 13, 2024) and Phase 2 (Prabowo's Transition Period: October 14, 2024 – April 25, 2025), allowing for comparative sentiment analysis during leadership changes.

For data labeling, a hybrid approach combined manual human annotation with LLM-assisted labeling. Two independent annotators manually labeled 30% of the dataset (9,451 tweets), resolving disagreements through structured guidelines to ensure consistency. The remaining 70% was labeled using a GPT-based LLM, leveraging zero-shot and few-shot classification for Bahasa Indonesia [18]. To validate LLM-generated labels, a random sample of 300 tweets was re-annotated by humans, with agreement measured via Cohen's Kappa (κ) and label-wise accuracy. Preprocessing involved tokenization, normalization (standardizing slang and casing), stop word removal, and stemming (reducing words like "pindah," "pemindahan," and "dipindahkan" to their root form) to prepare the data for feature extraction and modeling.

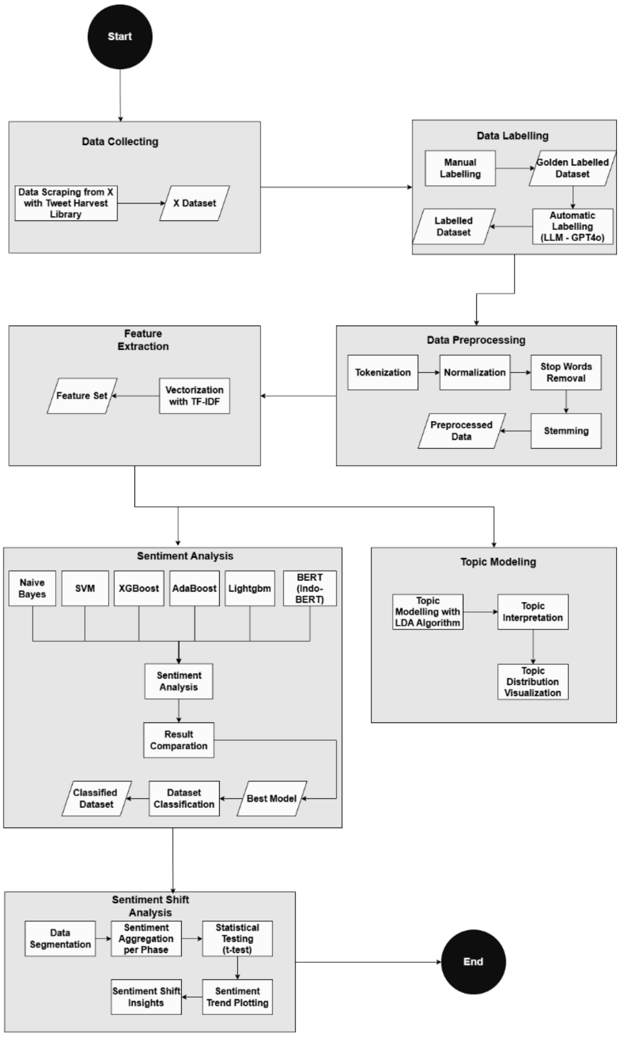

Figure 1. Research Methodology

The preprocessing phase ensures linguistic consistency and semantic focus, improving the accuracy of sentiment analysis and topic modeling. Feature extraction follows, where textual data is transformed into numerical formats using TF-IDF vectorization, which weights word importance relative to the corpus. N-grams (unigrams and bigrams) are incorporated to capture contextual phrases (e.g., "ibu kota"), enhancing feature representation for short texts like tweets. This structured data feeds into sentiment classification and topic modeling, enabling deeper analysis.

For sentiment analysis, multiple machine learning models—including Naïve Bayes, SVM, XGBoost, AdaBoost, and LightGBM—are trained on TF-IDF features to classify tweets as positive, negative, or neutral. Performance is evaluated using accuracy, precision, recall, and F1-score. Additionally, the IndoBERT transformer model is fine-tuned to handle nuanced language in social media text. This comparative approach identifies the most effective model for analyzing IKN-related sentiment while balancing computational efficiency. Sentiment shift analysis then examines changes between Jokowi's administration and Prabowo's transition period using t-tests and visualizations to detect statistically significant trends.

Finally, topic modeling via Latent Dirichlet Allocation (LDA) uncovers dominant themes in public discourse about IKN. The model, optimized using coherence scores, identifies key topics from TF-IDF vectorized n-grams. Stratifying results by sentiment reveals which issues are linked to positive (e.g., economic optimism) or negative (e.g., environmental concerns) reactions. Visualization tools like pyLDAvis aid interpretation, offering policymakers actionable insights into public priorities and concerns regarding Indonesia's new capital development.

RESULT AND DISCUSSION

This section presents the key findings of the research and analyzes their implications in the context of the study. It highlights the distribution of public sentiment, the dominant topics identified, and their relevance to addressing research questions while providing insights for policymakers and future research.

1. Data

The dataset used in this study was retrieved based on a curated set of keywords and hashtags commonly associated with the discourse surrounding Indonesia's Ibu Kota Nusantara (IKN) project. These keywords include popular terms such as #IKN, #IbuKotaBaru, IKN Indonesia, Pemindahan Ibu Kota, Ibu Kota Baru, Jokowi IKN, Prabowo IKN, Proyek IKN, Pembukaan IKN, and variations like #ibukotanusantara and ikn indonesia. These terms were selected following a comprehensive review of online conversations to ensure their relevance in capturing the diverse perspectives within the IKN discourse.

In total, 9,451 tweets were collected, spanning the period from April 9, 2017, to April 25, 2025, covering both the administration of President Joko Widodo and the transition period to President-elect Prabowo Subianto. This rich dataset supports both sentiment classification modeling and comparative temporal analysis, with keywords that span positive, negative, and neutral discussions surrounding the IKN policy.

Following data retrieval, all tweets underwent preprocessing, including cleaning, tokenization, normalization, stop-word removal, and stemming, to ensure data quality and consistency for downstream tasks. The preprocessed dataset was then annotated using a hybrid strategy: approximately 30% of the data was manually annotated by two independent human annotators, while the remaining 70% was labeled using a GPT-based Large Language Model (LLM) guided by structured prompts consistent with the manual annotation framework.

To evaluate the consistency between human and LLM-generated annotations, a random subset of 300 overlapping tweets was re-annotated manually. Among these, 57 tweets with completely matched sentiment labels were used to compute Cohen's Kappa, resulting in a score of 0.818—which, based on the Landis and Koch benchmark, reflects an "almost perfect agreement." This result validates the reliability of the LLM-generated annotations and confirms their suitability for training robust sentiment classification models. This result confirms that LLM-assisted annotations are highly consistent with human judgment, supporting their use as a scalable alternative to traditional annotation. Prior research also affirms that GPT-4 can perform sentiment labeling at a level comparable to or better than crowd workers, particularly in low-resource contexts [21].

In summary, the integration of LLMs with human oversight enabled the creation of a validated, scalable sentiment dataset. This method ensures that label quality remains within an acceptable margin of error while drastically reducing the manual effort required—offering both efficiency and accuracy for downstream model training.

The distribution of sentiments within the dataset, shown in Figure 3, reveals that most tweets are classified as neutral 5,997 tweets, while positive and negative sentiments account for 1,397 and 2,057 tweets, respectively. This sentiment distribution aligns with the nature of the IKN project—being a large-scale government initiative with both proponents and critics.

Moreover, the temporal span of the dataset, which covers multiple years and political phases, enhances the robustness of this study. By incorporating data from both the Jokowi administration and the transition period to Prabowo Subianto, the dataset enables a detailed examination of how public sentiment toward IKN has evolved across political leadership changes. This longitudinal perspective provides an empirical foundation for understanding the shifting dynamics of public opinion and the impact of political transitions on policy perception.

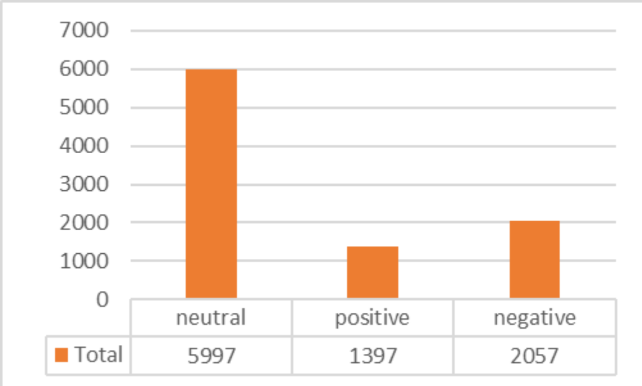

Figure 2. Distribution of X Data

2. Sentiment Analysis Results

To address Research Question 1 (RQ1), a comparative evaluation of various sentiment classification models was conducted, ranging from classical machine learning algorithms to state-of-the-art transformer-based models. Each model was assessed using four evaluation metrics: precision, recall, F1-score, and accuracy. The models were also tested across different n-gram configurations—unigram, bigram, and trigram—where applicable, to explore how context length affects model performance.

Among the traditional models, the Support Vector Machine (SVM) consistently achieved the highest performance. With unigram features, SVM reached an F1-score of 0.78 and an accuracy of 81.62%, surpassing other classical models such as Naïve Bayes, XGBoost, and LightGBM. Interestingly, SVM also performed robustly with bigram and trigram features, though there was a slight drop in performance compared to unigram, indicating that increasing n-gram size does not always result in performance gains.

Naïve Bayes, while efficient and lightweight, showed limitations in recall—particularly for bigram and trigram models—even though it scored high in precision (up to 0.86). This suggests that Naïve Bayes tends to over-predict the dominant class, making it less suitable for imbalanced or nuanced data distributions.

XGBoost and LightGBM demonstrated balanced results, with F1-scores ranging between 0.68 and 0.70, and accuracy between 72.97% and 76.22% across all n-gram variations. These models performed better than Adaboost, which recorded the lowest scores across all metrics, with a maximum accuracy of only 69.73%.

The best-performing model overall was BERT (IndoBERT), which significantly outperformed all other models with an F1-score of 0.81 and an accuracy of 85%. The transformer-based architecture of IndoBERT, pre-trained on large-scale Indonesian corpora, allows the model to capture deeper contextual and semantic relationships in the text, making it particularly effective for analyzing complex public discourse on social media platforms.

This comparative analysis demonstrates that simpler models like Naïve Bayes and Adaboost, while fast and easy to implement, sacrifice predictive performance, particularly on nuanced sentiment tasks. Meanwhile, more complex models like SVM and IndoBERT offer superior reliability and insight, though at a higher computational cost. As a result, IndoBERT was selected as the best model for further sentiment classification tasks in this study.

Table I. Comparative Analysis

Model	n-gram	precision	recall	f1-score	accuracy
adaboost	unigram	0.71	0.52	0.52	0.6757
adaboost	bigram	0.67	0.57	0.58	0.6973
adaboost	trigram	0.69	0.56	0.57	0.6973
lightgbm	unigram	0.69	0.67	0.68	0.7297
lightgbm	bigram	0.7	0.7	0.69	0.7351
lightgbm	trigram	0.68	0.65	0.66	0.7189
Naïve-Bayes	unigram	0.78	0.54	0.55	0.7027
Naïve-Bayes	bigram	0.84	0.53	0.53	0.7027
Naïve-Bayes	trigram	**0.86**	0.54	0.54	0.7135
svm	unigram	0.82	0.75	0.78	0.8162
svm	bigram	0.78	0.71	0.73	0.7838
svm	trigram	0.78	0.7	0.73	0.7784
xgboost	unigram	0.69	0.7	0.69	0.7351
xgboost	bigram	0.72	0.69	0.7	0.7514
xgboost	trigram	0.77	0.68	0.7	0.7622
BERT (Indo-BERT)		**0.83**	**0.79**	**0.81**	**0.85**

3. Sentiment Dynamics Across Political Phases

To examine how public sentiment toward IKN policies has changed over time, particularly across different political periods, the dataset was segmented into two phases: Phase 1 (Joko Widodo's Administration) and Phase 2 (Prabowo Subianto's Transition Period). This analysis answers Research Question 2 (RQ2), which investigates whether significant sentiment differences exist across the two leadership eras.

Using the sentiment classifications from the best-performing model (IndoBERT), tweets were first grouped based on their timestamp into the two phases. Sentiment polarity was quantified numerically as positive as +1, neutral as 0, and negative as -1 and a two-sample independent t-test was conducted to statistically compare the mean sentiment scores.

The t-test results are as follows T-statistic 3.94 and P-value: 8.28×10^{-5}. With a p-value well below the standard threshold of 0.05, the results indicate a statistically significant difference in sentiment between the two phases. This allows us to reject the null hypothesis and conclude that the shift in political leadership from Jokowi to Prabowo has had a measurable impact on how the public perceives IKN-related policies.

To further interpret the sentiment dynamics, time-series plots were generated for each phase:

Figure 1 displays the sentiment trend during Phase 1 (Jokowi's Administration). The chart reveals a steady increase in negative sentiment from 2017 to 2023, while positive sentiment fluctuates modestly, peaking in 2022 before dropping in 2023. Neutral sentiment remains consistently dominant throughout the phase, indicating a largely cautious or undecided public stance.

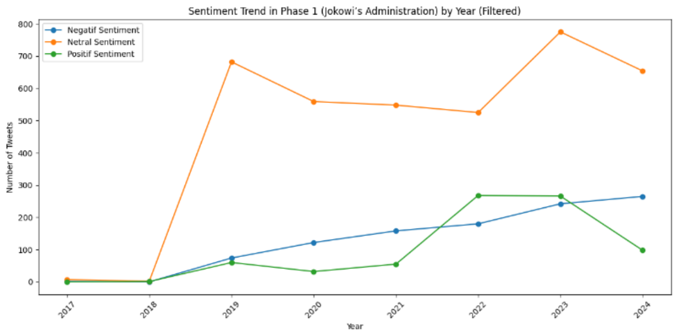

Figure 3. Sentiment Trend in Phase 1

Figure 3 shows sentiment dynamics during Phase 2 (Prabowo's Transition Period). Here, a sharp spike in neutral sentiment is observed in April 2025, coinciding with peak transition discourse and possibly media amplification. Interestingly, both negative and positive sentiments also increased in the same period, suggesting rising public polarization or increased engagement with the topic.

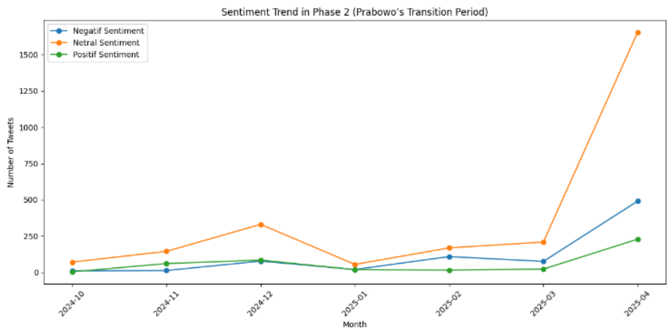

Figure 4. Sentiment Trend in Phase 2

These findings highlight how sentiment not only varies statistically between leadership periods but also responds dynamically to political context and external events. The data suggest that the announcement of Prabowo's presidency may have revitalized public discourse generating both optimism and criticism—thus reinforcing the need for sentiment-aware policy communication during transitional periods.

4. Topic Modeling Results and Interpretation

To address the third research question regarding the dominant issues and themes shaping public sentiment toward the IKN (Indonesia's new capital city) policy, we conducted topic modeling using the Latent Dirichlet Allocation (LDA) technique on Twitter data collected across two temporal phases: Phase 1 (Joko Widodo's Administration) and Phase 2 (Prabowo Subianto's Transition Period).

In Phase 1, which includes 5,569 tweets, the LDA model generated a coherence score of 0.4315, indicating moderate interpretability of the topic structure. As shown in Table II, the dominant keywords across the five extracted topics include "ikn," "kota" (city), "nusantara," "indonesia," "jokowi," "pindah" (relocation), "bangun" (build), and "negara" (nation). The frequent appearance of politically charged terms such as "jokowi," "prabowo," and "presiden" underscores the strong association between IKN and the broader political narrative. Additionally, economic terms like "apbn" (state budget) and "biaya" (cost) indicate that fiscal accountability was a recurrent theme of concern. The presence of English words such as "the," "of," and "capital" further suggests that the discourse was partly shaped by references to media coverage or international comparisons. The topic distribution shows a strong concentration in a single dominant topic, indicating that during this phase, public discussions were largely centered around the justification, symbolism, and political motivations underlying the relocation of the capital.

Table II. Topic Contribution Phase 1

Topic	**Terms**
Topic 0	ikn, jokowi, yg, kota, pindah, pembangunanikn, prabowo, iknnusantara, kepala
Topic 1	kota, nama, ibukota, jakarta, Indonesia, negara, desain, nusantara, ikn, yg
Topic 2	the, of, ikn, in, to, and, indonesia, is, for, ikon
Topic 3	Ikn, yg, tdk, dlm, utk, udah, amp. apbn, biaya, makan
Topic 4	ikn, kota, Nusantara, bangun, indonesia, negara, jokowi,pindah, presiden, kalimantan

The word cloud visualization for this phase (Figure 5) reinforces these findings by highlighting the visual prominence of "ikn," "kota," "pindah," and "jokowi" as the most frequently mentioned terms. This indicates that public discourse during this phase was heavily concentrated on symbolic identity, political leadership, and the rationale for the capital relocation, rather than implementation details or technical dimensions.

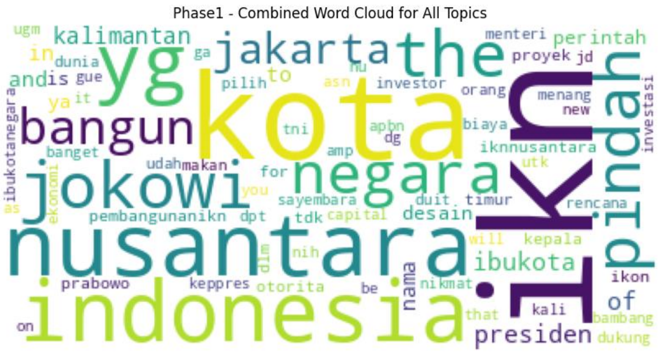

Figure 5.Word Cloud Visualization Phase 1

In terms of topic distribution (Figure 6), one topic significantly dominates the corpus, accounting for over 60% of the tweets. This asymmetrical distribution suggests that public sentiment in Phase 1 was narrowly focused on a single, overarching narrative, reflecting either intense support or scrutiny centered around the initial stages of the project announcement and justification. The limited diversity in topic engagement also suggests that, at this stage, the public discourse was less fragmented and more unified—albeit polarized—around a few core issues.

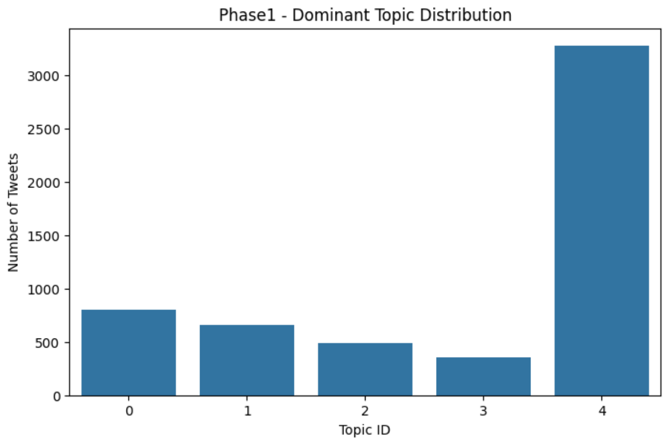

Figure 6. Dominant Topic Distribution Phase 1

In contrast, Phase 2, which contains 3,855 tweets, resulted in a higher coherence score of 0.4995, suggesting better topic clarity. The emerging themes in this phase reflect a more multifaceted public discourse. As shown in Table III terms like "ikn," "kota," "nusantara," and "bangun" remain prominent, while new words such as "asn" (civil servants), "otorita" (authority), "titik" (reference point), "tikus" (corruptor), and "smart city" indicate a shift in public attention from abstract planning to implementation and governance concerns. The use of symbolic language, such as "tikus," may reflect growing public cynicism or criticism toward perceived inefficiencies or corruption. Furthermore, the distribution of topics in this phase is significantly more balanced, with no single topic overwhelmingly dominating the discourse. This suggests a broader diversification of public interest, encompassing not only infrastructure development but also logistical challenges, bureaucratic readiness, and technological aspirations associated with the IKN project.

Table III. Topic Contribution Phase 2

Topic	Terms
Topic 0	ikn, asn, pindah, kota, tikus, negara, nusantara
Topic 1	Negara, Nusantara, aparatur, Prabowo, the, ikn, kota
Topic 2	Ikn, Indonesia, iknnusantara, Pembangunanikn, indonesiamaju, Indonesiaemas, maju, smartcity, perintah, kotamodern
Topic 3	Ikn, Jokowi, polri, Mahakam, poldakaltim, divhumas, coolingsystem, lebaran, kaltim, yg
Topic 4	Ikn, kota, Nusantara, bangun, otorita, Indonesia, titik, anggar, Prabowo, tugu

Overall, the topic modeling results reveal an evolution in societal concerns and expectations. In the early stages, public sentiment was largely shaped by the need to understand the rationale and symbolism of the capital relocation. However, as policy implementation progressed, discussions shifted toward concrete execution, institutional preparedness, and critical scrutiny. These findings highlight a transition from curiosity and symbolic support to more pragmatic concerns over governance, inclusiveness, and the real-world impact of IKN on everyday lives. Such insights emphasize the importance of adaptive and transparent policy communication that aligns with evolving public expectations.

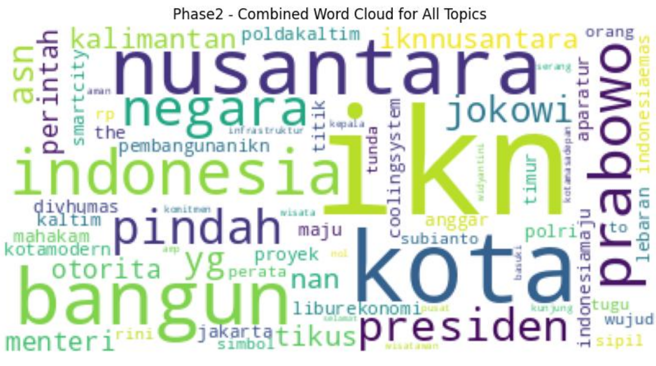

Figure 7. Word Cloud Visualization Phase 2

The combined word cloud for Phase 2 (Figure 7) reveals a broader lexical field than in Phase 1. In addition to foundational terms like “ikn” and “kota,” the prominence of terms such as “asn,” “tikus,” and “otorita” in the visualization illustrates heightened public attention to bureaucratic mobilization, institutional responsibility, and suspicions of corruption. This indicates that disillusionment or skepticism became more pronounced during the transition period, potentially fueled by media narratives and real-world observations of project execution.

Crucially, the topic distribution in Phase 2 (Figure 8) is substantially more balanced compared to Phase 1. All five topics appear with relatively even frequency, suggesting that public discourse became more fragmented and multifaceted. This balanced spread reflects the emergence of multiple coexisting narratives: some tweets celebrate IKN as a futuristic smart city, others criticize potential corruption (“tikus”), and still others focus on the logistical challenges faced by civil servants (“asn pindah”).

Such diversity implies that public engagement with IKN policies matured, transitioning from a unidimensional, politically symbolic focus to a multidimensional discourse that encompasses infrastructure, governance, fiscal discipline, administrative feasibility, and even environmental or technological aspects. This evolution reveals a more critical and informed public, demanding not just visionary promises, but also clear and credible implementation strategies.

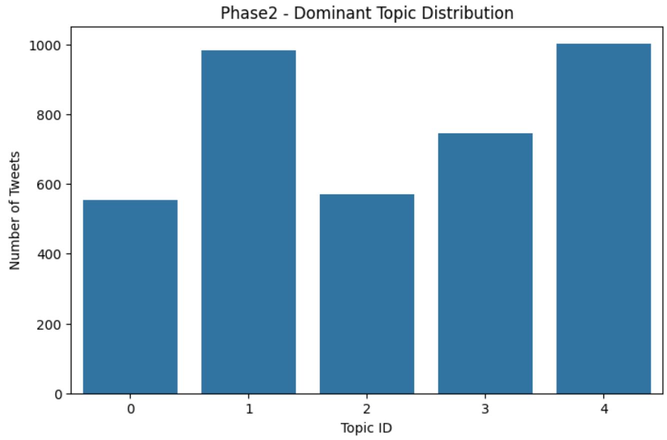

Figure 8. Dominant Topic Distribution Phase 1

5. Insights and Implications

The topic modeling results reveal a dynamic evolution in public discourse surrounding Indonesia’s IKN policies, reflecting the changing priorities, concerns, and expectations of the general public. During Phase 1, the discourse was heavily dominated by a few central themes—particularly the symbolic and political dimensions of relocating the capital. The dominance of terms such as “ikn,” “jokowi,” and “pindah” across both topic lists and the word cloud suggests that the conversation was driven by the novelty and nationalistic framing of the project. At this stage, IKN was still primarily perceived as an idea and an aspirational vision promoted by government leadership.

In contrast, Phase 2 presents a markedly different public narrative. The emergence of terms like “asn,” “otorita,” and “tikus” suggests growing public focus on implementation, institutional capacity, and concerns about governance integrity. This shift indicates that once the project moved from planning to execution, the public's expectations also shifted from symbolic alignment and political endorsement to functional delivery and accountability. Notably, the appearance of sarcasm-laden terms such as “tikus” reflects a deepening public cynicism, potentially fueled by previous experiences with government megaprojects, or perceptions of opacity in decision-making.

The word cloud visualizations support this interpretation. While the Phase 1 cloud displays concentrated and predictable terms around capital relocation and national identity, the Phase 2 cloud is broader and thematically varied—featuring terminology associated with administrative procedures ("asn"), infrastructure symbols ("titik," "tugu"), and future-oriented terms like "smart city." This indicates an increased diversification of public attention, wherein citizens are no longer just reacting to high-level announcements, but are also actively scrutinizing the operational and ethical dimensions of the policy.

The topic distribution charts offer another important signal. Phase 1 shows a clear dominance of one or two topics, suggesting that discourse was relatively uniform and potentially shaped by a top-down narrative from state officials and media. In Phase 2, however, the relatively even distribution of all five topics suggests a decentralization of discourse, where no single issue monopolizes public attention. This may indicate greater public agency and critical engagement, as citizens selectively engage with different dimensions of the project based on their interests, concerns, or proximity to impact.

From a policy standpoint, this shift has important implications. First, it underscores the need for phase-specific public engagement strategies. In the early stages of national projects, governments may be able to shape public perception through framing, vision statements, and symbolic gestures. However, as projects advance, expectations become more grounded, and communication must adapt to provide technical details, measurable progress, and accountability mechanisms.

Second, the rising presence of critical terms and diversified themes points to a growing demand for transparency, inclusion, and responsiveness. The government must go beyond broadcasting success stories; it must establish feedback loops where public input is acknowledged and integrated. For instance, criticism about bureaucratic readiness (e.g., through terms like "asn") signals the need to support affected civil servants through change management and clear communication. Meanwhile, discourse around potential corruption or inefficiency ("tikus") indicates that anti-corruption messaging and safeguards must be more visible and accessible to the public.

Third, the findings highlight the value of topic modeling as a real-time monitoring tool for policymaking. By identifying evolving public concerns and issue clusters, authorities can tailor interventions—be they public information campaigns, stakeholder consultations, or project revisions—to align with public sentiment. This aligns with global trends in evidence-based governance and the use of social media analytics to guide policy refinement.

Finally, the shift from political figures (e.g., "jokowi," "prabowo") to institutional and thematic terms (e.g., "otorita," "smartcity") suggests that public ownership of the discourse is increasing. As citizens begin to engage with the project at a functional and civic level, there is an opportunity to co-create solutions and build legitimacy for long-term success—provided that the state remains open and accountable in its approach.

CONCLUSION

This study investigated two key research questions: whether significant shifts in public sentiment occurred across different political phases of Indonesia's IKN project, and which methods were most effective for sentiment analysis and thematic extraction. The sentiment shift analysis (RQ1) revealed statistically significant differences in public sentiment between Joko Widodo's administration and Prabowo Subianto's transition period, as confirmed by t-tests, demonstrating how leadership changes and policy announcements influence public perception of major national projects. For RQ2, comparative evaluation of sentiment analysis methods showed that IndoBERT outperformed traditional machine learning models (Naïve Bayes, SVM, AdaBoost, XGBoost, LightGBM) in accuracy and nuance, though ensemble methods remained valuable for computational efficiency. Topic modeling identified recurring themes—environmental sustainability, socio-economic concerns, infrastructure development, governance quality, and transparency—that shaped public discourse, offering policymakers actionable insights. By combining advanced machine learning with statistical rigor, this study provides a comprehensive understanding of IKN-related public sentiment dynamics, suggesting future research could expand to additional platforms, demographic factors, and real-time communication impacts to further enhance policy responsiveness.

REFERENCES

[1] Nurkaidah, A. Anas, and T. Baharuddin, "Implementation of environmental policies on the development of a new capital city in Indonesia," *Cogent Soc. Sci.*, vol. 10, no. 1, Dec. 2024, doi: 10.1080/23311886.2023.2297764.

[2] M. Nokkaew *et al.*, "Analyzing online public opinion on Thailand-China high-speed train and Laos-China railway mega-projects using advanced machine learning for sentiment analysis," *Soc. Netw. Anal. Min.*, vol. 14, no. 1, Dec. 2023, doi: 10.1007/s13278-023-01168-8.

[3] A. Sharma and H. Shekhar, "Intelligent Learning based Opinion Mining Model for Governmental Decision Making," *Procedia Comput. Sci.*, vol. 173, pp. 216–224, 2020, doi: 10.1016/j.procs.2020.06.026.

[4] Z. A. Faridzi, D. Pramesti, and R. Y. Fa'rifah, "A Comparison of Oversampling and Undersampling Methods in Sentiment Analysis Regarding Indonesia Fuel Price Increase Using Support Vector Machine," in *2023 International Conference on Advancement in Data Science, E-learning and Information System (ICADEIS)*, Bali, Indonesia: IEEE, Aug. 2023, pp. 1–6. doi: 10.1109/icadeis58666.2023.10270851.

[5] V. A. Ayma Quirita, J. David Cárdenas, W. Aliaga, V. H. Ayma Quirita, A. Palacios, and R. B. Sierra, "Peruvian Presidential Debates in the Elections of 2021 in Twitter/X: A Sentiment Analysis Approach," *IEEE Access*, vol. 12, pp. 138386–138398, 2024, doi: 10.1109/access.2024.3465503.

[6] D. S. Zainulabdeen, M. Çevik, and M. M. Abdulrazzaq, "A Comparative Study of Classification Algorithms for Sentiment Analysis of Covid-19 Vaccine Opinions Using Machine Learning," in *2024 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*, Istanbul, Turkiye: IEEE, May 2024, pp. 1–6. doi: 10.1109/hora61326.2024.10550643.

[7] A. A. Dwisanny and S. Supatmi, "Twitter Opinion Sentiment Analysis Based on New Student Admission Zoning Issues Using the Naïve Bayes and TensorFlow Methods," in *2023 9th International Conference on Signal Processing and Intelligent Systems (ICSPIS)*, Bali, Indonesia: IEEE, Dec. 2023, pp. 1–8. doi: 10.1109/icspis59665.2023.10402745.

[8] M. A. Khadija, I. S. D. Jayanti, and F. U. Nimah, "Towards Smart City: Aspect Based Sentiment Analysis of Indonesian Public Aspiration Complaints Data Using Machine Learning," in *2024 7th International Conference on Informatics and Computational Sciences (ICICoS)*, Semarang, Indonesia: IEEE, Jul. 2024, pp. 215–220. doi: 10.1109/icicos62600.2024.10636859.

[9] Kartini, A. K. Darmawan, R. I. Syah, and M. Makruf, "Sentiment Analysis of Social Media for Indonesian m-Health PeduliLindungi Mobile-Apps(PLMA) with Lexicon-Based and Support Vector Machine Approach," in *2023 IEEE 9th Information Technology International Seminar (ITIS)*, Batu Malang, Indonesia: IEEE, Oct. 2023, pp. 1–7. doi: 10.1109/itis59651.2023.10419994.

[10] A. Alotaibi and F. Nadeem, "Leveraging Social Media and Deep Learning for Sentiment Analysis for Smart Governance: A Case Study of Public Reactions to Educational Reforms in Saudi Arabia," *Computers*, vol. 13, no. 11, p. 280, Oct. 2024, doi: 10.3390/computers13110280.

[11] M. Chen, H. Xu, Y. Wu, and J. Wu, "Sentiment Analysis of Hotel Reviews based on BERT and XGBoost," in *2024 3rd International Conference on Computer Technologies (ICCTech)*, Bali, Indonesia: IEEE, Feb. 2024, pp. 11–15. doi: 10.1109/icctech61708.2024.00011.

[12] Y. Shen and P. K. Zhang, "Financial Sentiment Analysis on News and Reports Using Large Language Models and FinBERT," in *2024 IEEE 6th International Conference on Power, Intelligent Computing and Systems (ICPICS)*, Shenyang, China: IEEE, Jul. 2024, pp. 717–721. doi: 10.1109/icpics62053.2024.10796670.

[13] R. Jaiswal, S. Gupta, and A. K. Tiwari, "Decoding mood of the Twitterverse on ESG investing: opinion mining and key themes using machine learning," *Manag. Res. Rev.*, vol. 47, no. 8, pp. 1221–1252, Jul. 2024, doi: 10.1108/mrr-07-2023-0526.

[14] M. M. Alexandrovna and M. A. Mohsin, "Tweet Sentiment Analysis with CNN and XG-BOOST," vol. 21, no. 4, 2023.

[15] Z. Wang, "Sentiment Analysis of Mobile Phone Reviews Using XGBoost and Word Vectors," *ITM Web Conf.*, vol. 70, p. 03018, 2025, doi: 10.1051/itmconf/20257003018.

[16] L. Jixian, A. Gang, S. Zhihao, and S. Xiaoqiang, "Social Media Multimodal Information Analysis based on the BiLSTM-Attention-CNN-XGBoost Ensemble Neural Network," *Int. J. Adv. Comput. Sci. Appl.*, vol. 13, no. 12, 2022, doi: 10.14569/ijacsa.2022.0131215.

[17] A. Onan, "Sentiment analysis on product reviews based on weighted word embeddings and deep neural networks," *Concurr. Comput. Pract. Exp.*, vol. 33, no. 23, Dec. 2021, doi: 10.1002/cpe.5909.

[18] Hidayaturrahman and I. Prawira, "Leveraging Zero-Shot Learning in Large Language Models for Sentiment Analysis: A Comparative Study on the Indonesian Language," in *2024 International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS)*, Jakarta, Indonesia: IEEE, Nov. 2024, pp. 614–619. doi: 10.1109/icimcis63449.2024.10956237.

[19] A. Ross, V. L. Willson, A. Ross, and V. L. Willson, "Independent samples T-test," *Basic Adv. Stat. Tests Writ. Results Sect. Creat. Tables Fig.*, pp. 13–16, 2017.

[20] H. Jang, E. Rempel, D. Roth, G. Carenini, and N. Z. Janjua, "Tracking COVID-19 Discourse on Twitter in North America: Infodemiology Study Using Topic Modeling and Aspect-Based Sentiment Analysis," *J Med Internet Res*, vol. 23, no. 2, p. e25431, Feb. 2021, doi: 10.2196/25431.

[21] A. H. Nasution and A. Onan, "ChatGPT Label: Comparing the Quality of Human-Generated and LLM-Generated Annotations in Low-Resource Language NLP Tasks," *IEEE Access*, vol. 12, pp. 71876–71900, 2024, doi: 10.1109/access.2024.3402809.