Journal of Information System Research (JOSH)
Volume 7, No. 2, Januari 2026, pp 349−359
ISSN 2686-228X (media online)
https://ejurnal.seminar-id.com/index.php/josh/
DOI 10.47065/josh.v7i2.8839

Sentiment Analysis on X, TikTok, and Instagram on Indonesian Capital relocation using Support Vector Machine

Syawalian Rais Dwi Jayanto*, Suprihadi

Faculty of Information Technology, Informathic Engineering, Satya Wacana Christian University, Salatiga
Jl. Diponegoro No. 52-60, Salatiga, Sidorejo District, Salatiga City, Central Java, Indonesia
Email: [1,*]672022198@student.uksw.edu, [2]suprihadi@uksw.edu
Correspondence Author Email: 672022198@student.uksw.edu
Submitted: **04/12/2025**; Accepted: **06/01/2026**; Published: **23/01/2026**

Abstract−This study examines public sentiment toward Indonesia's new capital city, Ibu Kota Nusantara (IKN), across three major social media platforms: X, TikTok, and Instagram. The research aims to identify how public perceptions differ across platforms and to understand their implications for policy communication. A total of approximately 6,000 user comments collected up to March 2025 were processed through standard text-mining procedures, including cleaning, tokenization, stop-word removal, and stemming. The text data were converted into numerical features using the Term Frequency–Inverse Document Frequency (TF-IDF) technique and classified using a linear Support Vector Machine (SVM) model. Model evaluation with a 20% hold-out test set yielded an accuracy of 90.23% and a macro F1-score of 0.8905. The analysis shows that overall sentiment toward IKN is predominantly positive, with Instagram and TikTok generating more supportive narratives, while X displays a higher concentration of critical or negative comments. These findings highlight significant platform-specific differences that can inform more effective public communication strategies regarding the IKN project.

Keywords: Nusantara Capital City; Sentiment Analysis; Social Media; Support Vector Machine; TF-IDF.

1. INTRODUCTION

The relocation of the Indonesian capital has generated extensive public discussion across various social media platforms. These discussions reflect diverse public opinions, including support, criticism, and neutrality toward the policy. Social media platforms such as X, TikTok, and Instagram have become major channels for expressing public sentiment due to their wide user base and real-time information dissemination. Therefore, sentiment analysis is necessary to understand public perception of the Indonesian capital relocation issue. This study applies a machine learning approach using the Support Vector Machine algorithm to classify public sentiment into positive, negative, and neutral categories based on text data collected from X, TikTok, and Instagram.

The relocation of the Nusantara Capital City (IKN) to East Kalimantan is a strategic national issue that has triggered diverse public opinions. These opinions are widely disseminated through three major social media platforms X (formerly Twitter), TikTok, and Instagram each with distinct communication characteristics and user demographics. Data from the official APJII report in 2024 indicate that TikTok users in Indonesia reached 157.6 million, Instagram 90.18 million, and X 25 million [1]. Law Number 3 of 2022 officially stipulates the relocation of IKN to East Kalimantan. Despite being enacted in 2022, the issue remains a matter of public concern. The planned relocation of civil servants (ASN), initially scheduled for September 2024, was postponed until 2025, reflecting ongoing dynamics [2]. Furthermore, diverse public criticisms and debates continue to emerge, ranging from concerns over the high budget to uncertainties about the readiness of infrastructure [3]. This indicates that the policy remains controversial even after being legally established. Recent studies indicate that TikTok plays a crucial role in shaping public sentiment toward the development of the Nusantara Capital City, as its short-form and viral content encourages rapid emotional engagement among users [14].

The urgency of this study lies in the need to comprehensively understand public opinion as part of the government's communication strategy. The IKN communication team stated that until 2045, their focus is to enhance public understanding of the relocation plan, manage the evolving narratives in society, and build public participation in the transition process [3]. Thus, mapping public sentiment becomes crucial as a foundation for determining appropriate communication strategies across digital channels. The content character of each platform influences how people express opinions. Instagram, which primarily emphasizes visuals, conveys opinions through photos and videos, with comment sections often reflecting public perspectives on certain issues [14]. TikTok uses short videos that quickly go viral to engage younger generations, and its rapidly spreading content has been utilized by politicians for campaigning [14]. Although X has fewer users, it is widely used for discussions on political matters and current issues. In the "algorithmic era," whoever raises an issue quickly and massively on social media can influence public opinion. These three platforms were chosen to evaluate emotions regarding IKN relocation because Instagram facilitates visual expression and public comments, TikTok promotes youth perspectives through creative videos, and X contributes to rapid discussions on national issues [4, 2].

The Support Vector Machine (SVM) algorithm was selected because it consistently demonstrates strong performance in text classification tasks, especially in distinguishing sentiment categories within complex and noisy social media data. Similar results were also reported in another study analyzing public sentiment on the IKN relocation issue using Twitter data, where Support Vector Machine demonstrated superior performance compared to probabilistic classifiers [12]. Previous research comparing Naive Bayes and Support Vector Machine for

Journal of Information System Research (JOSH)
Volume 7, No. 2, Januari 2026, pp 349−359
ISSN 2686-228X (media online)
https://ejurnal.seminar-id.com/index.php/josh/
DOI 10.47065/josh.v7i2.8839

sentiment analysis on Twitter data related to the Indonesian capital relocation shows that SVM consistently outperforms Naive Bayes in terms of classification accuracy [13]. Similar results were also reported in another study analyzing public sentiment on the IKN relocation issue using Twitter data, where Support Vector Machine demonstrated superior performance compared to probabilistic classifiers [12]. Prior research supports the effectiveness of SVM for public policy sentiment analysis. A study on public responses to the IKN relocation reported that a linear-kernel SVM achieved an accuracy of 92.70%, which further increased to 100% when using Stratified k-Fold validation, indicating high model stability and generalizability [14]. Other studies also reinforce this approach: sentiment analysis of public reactions to COVID-19 policies showed that linear SVM outperformed Naïve Bayes and Random Forest, achieving F1-scores above 0.88; research on political discourse on Twitter similarly found that SVM produced the most reliable classification results compared to KNN and Decision Tree; and an analysis of transportation policy debates demonstrated that SVM maintained high accuracy even with unbalanced datasets. Collectively, these findings confirm that SVM is a robust and suitable method for analyzing public sentiment toward the Indonesian capital relocation across X, TikTok, and Instagram [12].

This research focuses on two main questions. First, how are public sentiments distributed and compared regarding the IKN relocation issue across X, TikTok, and Instagram? Second, how effective is the SVM algorithm in classifying positive and negative sentiments across platforms, considering differences in data formats and user interactions that may affect classification consistency? The aim of this study is to analyze and compare the distribution of public sentiments regarding the IKN relocation on X, TikTok, and Instagram, as well as to evaluate the effectiveness of the SVM algorithm in classifying such sentiments.

Public policy issues has grown significantly in recent years, particularly through the utilization of social media data. Putra and Wibowo [13] examined public sentiment toward COVID-19 response policies in Indonesia using the Support Vector Machine (SVM) method with TF-IDF as the feature extraction technique. Their research procedure included collecting Twitter data via API, text cleaning, manual labeling, and model evaluation, which yielded an accuracy of 88%. The findings revealed that negative sentiment dominated during the early period of the pandemic, influenced by high transmission rates and policy uncertainty.

Research on sentiment analysis related to ctively in classifying public opinion on policy-related issues. Furthermore, Saputra et al. [14] analyzed public perceptions of national infrastructure development through machine learning–based sentiment analysis. They utilized public comments from Instagram and YouTube, employing data crawling, preprocessing, TF-IDF, and classification using a combination of SVM and Naïve Bayes. Their research procedures indicated that SVM produced higher accuracy than Naïve Bayes, reaching 91.2%. Their findings highlight the distinct characteristics of each platform in shaping public opinion, particularly between visually oriented and text-based media.

Another study by Ramadhani and Gunawan [12] investigated public responses to energy subsidy policies using TikTok data. The methods applied included scraping video comments, text preprocessing, double annotator labeling, and analysis using SVM and Random Forest. SVM achieved the highest performance with an F1-score of 0.89. This study underscores that TikTok exhibits rapid and emotionally driven opinion dissemination patterns, contributing to unique sentiment dynamics compared to other platforms.

Meanwhile, Lestari and Prabowo [13] assessed public sentiment toward corruption issues using data from X (formerly Twitter). They employed both lexicon-based and supervised learning approaches to compare their effectiveness. The research procedures included tokenization, stopword removal, stemming, and n-gram modeling. The findings showed that the SVM model with bigram features yielded more accurate classifications than lexicon-based approaches, particularly for Indonesian texts with complex morphological variations.

Specifically regarding the Nusantara Capital City (IKN) issue, several studies have been conducted, although still limited in number. For example, Abdullah et al. [14] evaluated public sentiment toward the planned relocation of IKN using Twitter data. The study employed text mining, manual labeling, TF-IDF, and SVM. The results indicated the dominance of negative sentiment influenced by environmental concerns, budget issues, and infrastructure readiness. Although focused on a single platform, this study provides an initial insight into public perception patterns regarding IKN.

The object of this research is public opinion related to the issue of Indonesian capital relocation expressed on social media platforms X, TikTok, and Instagram. The data used in this study consist of text comments and posts collected using keyword-based crawling techniques related to the capital relocation issue. Data were collected within a specific time period to capture recent public sentiment. Each data instance was manually labeled into three sentiment classes: positive, negative, and neutral. The dataset includes a total number of samples that represent balanced and imbalanced sentiment distributions across the three platforms.

This study offers novelty compared to previous research by integrating cross-platform analysis from three major social media channels X, TikTok, and Instagram which exhibit distinct user characteristics and demographics. Unlike earlier studies that focused solely on a single platform, this research processes data separately for each medium to capture more accurate variations in discourse dynamics. Methodologically, the study employs a combination of unigram bigram TF-IDF and linear SVM optimized via grid search, yielding higher accuracy compared to simpler models used in prior studies. The research procedures are also more comprehensive, involving structured manual labeling, evaluation using platform-specific confusion matrices, and comparative analysis of model performance. The novelty of this study lies in its multi-platform approach, separated

Journal of Information System Research (JOSH)
Volume 7, No. 2, Januari 2026, pp 349−359
ISSN 2686-228X (media online)
https://ejurnal.seminar-id.com/index.php/josh/
DOI 10.47065/josh.v7i2.8839

data workflows for each medium, and in-depth exploration of inter-platform differences in public perception as a foundation for designing more effective policy communication strategies.

2. RESEARCH METHODOLOGY

The research stages employed in this study are structured based on the research workflow illustrated in Figure 1.

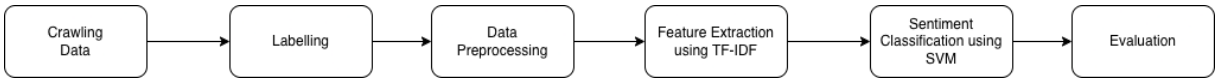

Figure 1. Research Method Flow

The research methodology in this study consists of several sequential stages. The first stage is data crawling, in which public comments related to the Indonesian capital relocation issue are collected from three social media platforms, namely X, TikTok, and Instagram. Data were obtained using keyword-based crawling techniques and collected up to March 2025, resulting in approximately 2,000 comments from each platform or around 6,000 comments in total.

The second stage is data labelling, where each collected comment is manually annotated into three sentiment categories: positive, negative, and neutral. Manual labelling was conducted to ensure labeling accuracy and to provide reliable ground truth data for supervised learning.

The third stage is data preprocessing, which aims to clean and normalize the raw text data. This stage includes data cleaning to remove URLs, mentions, hashtags, emojis, and non-alphabetic characters, followed by case folding to convert all text into lowercase. Tokenization is then performed to split text into individual word units, stopword removal is applied using the Indonesian stopword list, and stemming is conducted using the Sastrawi algorithm to convert words into their root forms.

The fourth stage is feature extraction, where the preprocessed text data are transformed into numerical representations using the Term Frequency–Inverse Document Frequency method with unigram and bigram features. This process assigns weights to words based on their frequency and importance within the dataset. Support Vector Machine was selected in this study because previous research confirms its stable and superior performance compared to Logistic Regression and Naïve Bayes in sentiment classification tasks on social media data [16]. The fifth stage is sentiment classification, where the extracted features are classified using the Support Vector Machine algorithm with a linear kernel. The dataset is divided into training and testing data using an 80:20 split, and hyperparameter tuning is conducted using grid search to determine the optimal regularization parameter. The combination of TF-IDF feature weighting and Support Vector Machine classification is widely adopted in sentiment analysis studies on the IKN issue due to its robustness in handling textual variation and class imbalance [17]. The final stage is model evaluation, which assesses the performance of the classification model using accuracy, precision, recall, and F1-score metrics. Confusion matrices are also generated for each platform to analyze misclassification patterns and compare sentiment classification performance across X, TikTok, and Instagram.

3. RESULT AND DISCUSSION

3.1 Crawling Data Results

This study collected public comments related to the Indonesian capital relocation issue from three major social media platforms, namely X, TikTok, and Instagram, which serve as important online spaces for public discussion and opinion expression regarding national development policies. These platforms were selected because of their high user engagement and their ability to represent diverse demographic backgrounds and communication styles. Data collection was conducted using automated crawling and scraping techniques to obtain large-scale textual data in a systematic and efficient manner. Specifically, Tweet Harvest was used to extract comments and replies from X, Apify was employed to retrieve user comments from TikTok videos discussing the capital relocation issue, and IG Comment Export tools were utilized to collect comment data from relevant Instagram posts. Through this process, a total of approximately 6,000 public comments were successfully gathered, with around 2,000 comments collected from each platform, ensuring balanced data distribution and reducing platform-specific bias in the dataset.

The collected comments reflect a wide range of public opinions, attitudes, and emotional responses toward the Indonesian capital relocation policy, expressed through informal and unstructured language typical of social media communication. Proper preprocessing is essential to enhance data consistency, reduce ambiguity, and improve the performance of machine learning models used in this study. The results of the data crawling process from Instagram, X, and TikTok are presented in Figure 2, Figure 3, and Figure 4, respectively, to provide a visual overview of the volume and distribution of comments collected from each social media platform.

Journal of Information System Research (JOSH)
Volume 7, No. 2, Januari 2026, pp 349−359
ISSN 2686-228X (media online)
https://ejurnal.seminar-id.com/index.php/josh/
DOI 10.47065/josh.v7i2.8839

	platform	text
0	instagram	@dimasprasetyo18juli lo aja tanem diri dalam t...
1	instagram	NaN
2	instagram	buat orang2 yg ga pro ikn, percayalah suatu sa...
3	instagram	@deni_sup09 nandur o Dewe kebutuhan mu
4	instagram	Pak kalau butuh tukang AC atau tukang listrik ...
5	instagram	Semangat.terus.paak terimakasih kepada bapak j...
6	instagram	@adisk2912 iyow... Tak wa ne bapak e, sing tenang
7	instagram	Terbaik pa
8	instagram	Sehat selalu pak Jokowi..... JBU
9	instagram	Semoga di kota baru nanti gak ada yg buang sam...
10	instagram	@ahmadfaruqalkaffi siap tok lek kene iki
11	instagram	Yth bpk @jokowi tolong sejahterakan pekerja AS...
12	instagram	❤❤❤
13	instagram	@nabiel1984 jangan fitnah liat noh videonya an...
14	instagram	Lompatan ? Sepertinya pada moment tertentu saj...

Figure 2. Crawling Data Results from Instagram

	platform	full_text
0	tiktok	NaN
1	tiktok	❤❤❤❤❤
2	tiktok	😂
3	tiktok	😂😂😂😂😂
4	tiktok	😂😂😂
5	tiktok	hmmmmmmm
6	tiktok	cot cot
7	tiktok	enteng banget tu rahang.😡😡😡
8	tiktok	proyek Mulyono bkn Jokowi
9	tiktok	keputusan rakyat mana woi... 😂😂😂😂
10	tiktok	ijin DPR cuma formalitas..\nResminya ambisi pr...
11	tiktok	2 kata lucu : keputusan rakyat
12	tiktok	mulai cuci tangan, muali panik
13	tiktok	masyarakat mana oi😎 orang tbtb ada pemutusan p...
14	tiktok	hadehh mulyono rakyat yang mana noo

Figure 3. Crawling Data Results from TikTok

	platform	full_text
0	x	@O5Bravo Pemindahan Ibukota itu ok tapi kalau ...
1	x	Gubernur Kaltim : IKN akan Beroperasi Paling L...
2	x	Makanya mata rantainya diputus dengan pemindah...
3	x	@zhil_arf Barangkali itu salah satu motivasi p...
4	x	Waisak 2025 di Kaltim: Doa Damai untuk Dunia d...
5	x	@zhambernyowo @sarah_pndjtn Jadi??? Yg km 50 d...
6	x	Pemindahan IKN merupakan wujud nyata pemerinta...
7	x	Pemutarbalikan Kebijakan Populer Memutarbalikk...
8	x	Opini saya lainnya terbit di Koran terpopuler ...
9	x	Komisi II DPR RI : Kaji Kesiapan Infrastruktur...
10	x	OIKN: Hunian ASN di IKN Siap Tampung Pemindaha...
11	x	@tham878 Proyek negara di jadikan lahan binis ...
12	x	OIKN Jamin Hunian di IKN Bisa Tampung Pemindah...
13	x	OIKN: Hunian ASN di IKN siap tampung pemindaha...
14	x	@HarryPattyRM00 @Lucy_888888 @DokterTifa banya...

Figure 4. Crawling Data Results from X

3.2 Data Preprocessing Results

The data preprocessing stage was conducted to reduce noise and standardize textual data obtained from social media platforms. This stage is essential to ensure that the text data are suitable for feature extraction and sentiment classification. The preprocessing process consisted of data cleaning, case folding, tokenization, stopword removal, and stemming.

3.2.1 Data Cleaning

The data cleaning process aimed to remove irrelevant elements from the raw text, such as URLs, user mentions, hashtags, emojis, numbers, and special characters that do not directly contribute to sentiment interpretation. By removing these unnecessary elements, the cleaning process ensured that only meaningful and relevant textual information remained, thereby improving data quality and facilitating more accurate. An example of the data cleaning results for Instagram, X, and TikTok comments is shown in Figures 5, 6, and 7.

Journal of Information System Research (JOSH)
Volume 7, No. 2, Januari 2026, pp 349−359
ISSN 2686-228X (media online)
https://ejurnal.seminar-id.com/index.php/josh/
DOI 10.47065/josh.v7i2.8839

index	text	cleaned
0	@dimasprasetyo18juli lo aja tanem diri dalam tanah	lo aja tanem diri dalam tanah
1		
2	buat orang2 yg ga pro ikn, percayalah suatu saat di masa depan, kalian bahkan anak cucu kalian bangga mempunyai ibukota seperti ikn🙌	buat orang2 yg ga pro ikn percayalah suatu saat di masa depan kalian bahkan anak cucu kalian bangga mempunyai ibukota seperti ikn
3	@deni_sup09 nandur o Dewe kebutuhan mu	nandur o Dewe kebutuhan mu
4	Pak kalau butuh tukang AC atau tukang listrik di sana, kabar2 gih pak..🙏🙏😊	Pak kalau butuh tukang AC atau tukang listrik di sana kabar2 gih pak
5	Semangat.terus.paak terimakasih kepada bapak jokowi telah.memajukan.indonesia.	Semangatteruspaak terimakasih kepada bapak jokowi telahmemajukanindonesia
6	@adisk2912 iyow... Tak wa ne bapak e, sing tenang	iyow Tak wa ne bapak e sing tenang
7	Terbaik pa	Terbaik pa
8	Sehat selalu pak Jokowi..... JBU	Sehat selalu pak Jokowi JBU
9	Semoga di kota baru nanti gak ada yg buang sampah sembarangan	Semoga di kota baru nanti gak ada yg buang sampah sembarangan

Figure 5. Data Cleaning Results of Instagram

1 to 10 of 10 entries Filter

index	full_text	cleaned
0	@O5Bravo Pemindahan Ibukota itu ok tapi kalau model kaya IKN. Iya lucu saja. Alasannya tiap waktu saja semakin konyol. Semakin tidak waras saja orang yang bela IKN. Saya voter 03. Tapi anggap mereka orang gila saja dukung proyek mangkrak. Sudahlah mangkrak. Biaya sudah habis banyak.	Pemindahan Ibukota itu ok tapi kalau model kaya IKN Iya lucu saja Alasannya tiap waktu saja semakin konyol Semakin tidak waras saja orang yang bela IKN Saya voter 03 Tapi anggap mereka orang gila saja dukung proyek mangkrak Sudahlah mangkrak Biaya sudah habis banyak
1	Gubernur Kaltim : IKN akan Beroperasi Paling Lambat 2028 ... Pak @prabowo kenapa belum juga teken perpres pemindahan ASN ke IKN? https://t.co/Nn7dOWcnzv	Gubernur Kaltim IKN akan Beroperasi Paling Lambat 2028 Pak kenapa belum juga teken perpres pemindahan ASN ke IKN
2	Makanya mata rantainya diputus dengan pemindahan ibukota. Gara gara itu Kelas liberal dan Islamis modern benci banget sama IKN. Padahal mereka ini minoritas. Di pilpres aja pemilih capres pro IKN jumlahnya 58+16 persen kok. https://t.co/1TwHrYCjGO	Makanya mata rantainya diputus dengan pemindahan ibukota Gara gara itu Kelas liberal dan Islamis modern benci banget sama IKN Padahal mereka ini minoritas Di pilpres aja pemilih capres pro IKN jumlahnya 5816 persen kok
3	@zhil_arf Barangkali itu salah satu motivasi pemindahan IKN ke Kaltim krn di sana kaum progresif liberal yg menamakan diri civil society maupun Islamis garis keras tidak terlalu kuat.	Barangkali itu salah satu motivasi pemindahan IKN ke Kaltim krn di sana kaum progresif liberal yg menamakan diri civil society maupun Islamis garis keras tidak terlalu kuat
4	Waisak 2025 di Kaltim: Doa Damai untuk Dunia dan Kelancaran Proses Pemindahan IKN https://t.co/ERxWDON1Fp	Waisak 2025 di Kaltim Doa Damai untuk Dunia dan Kelancaran Proses Pemindahan IKN
5	@zhambernyowo @sarah_pndjtn Jadi??? Yg km 50 dia presiden nya. Yg janji pesan mobil dia. Yg ambisi pemindahan IKN dia. Yg mengatakan IPK kurang 2 dia. Pemuja nya keras keras ga menentu. Yg di puja lembek lembek bau. Zholim....	Jadi Yg km 50 dia presiden nya Yg janji pesan mobil dia Yg ambisi pemindahan IKN dia Yg mengatakan IPK kurang 2 dia Pemuja nya keras keras ga menentu Yg di puja lembek lembek bau Zholim
6	Pemindahan IKN merupakan wujud nyata pemerintah dalam meningkatkan kesejahteraan dan mempererat persatuan seluruh masyarakat Indonesia. https://t.co/YGPJ17d0sN	Pemindahan IKN merupakan wujud nyata pemerintah dalam meningkatkan kesejahteraan dan mempererat persatuan seluruh masyarakat Indonesia
7	Pemutarbalikan Kebijakan Populer Memutarbalikkan pencapaian seperti pemindahan IKN dan pembangunan infrastruktur menyebutnya sebagai pemborosan atau upaya pencitraan untuk mengurangi dukungan publik.	Pemutarbalikan Kebijakan Populer Memutarbalikkan pencapaian seperti pemindahan IKN dan pembangunan infrastruktur menyebutnya sebagai pemborosan atau upaya pencitraan untuk mengurangi dukungan publik
8	Opini saya lainnya terbit di Koran terpopuler di Jawa Tengah Suara Merdeka edisi 8/5/2025 terkait Karut Marut Pemindahan Aparatur Negara ke IKN. Dapat diakses melalui link berikut https://t.co/6oNB0sAQ23. #NicholasMartuaSiagian #Koran #SuaraMerdeka https://t.co/RJlWAzSiGR	Opini saya lainnya terbit di Koran terpopuler di Jawa Tengah Suara Merdeka edisi 852025 terkait Karut Marut Pemindahan Aparatur Negara ke IKN Dapat diakses melalui link berikut
9	Komisi II DPR RI : Kaji Kesiapan Infrastruktur Terkait Pemindahan ASN Ke IKN - TVR 120 https://t.co/rFCyMVTcAM	Komisi II DPR RI Kaji Kesiapan Infrastruktur Terkait Pemindahan ASN Ke IKN TVR 120

Figure 6. Data Cleaning Results of X

1 to 10 of 10 entries Filter

index	full_text	cleaned
0		
1	❤️❤️❤️❤️❤️	
2	😊	
3	😂😂😂😂😂	
4	😂😂😂	
5	hmmmmmmm	hmmmmmmm
6	cot cot	cot cot
7	enteng banget tu rahang.😡😡😡	enteng banget tu rahang
8	proyek Mulyono bkn Jokowi	proyek Mulyono bkn Jokowi
9	keputusan rakyat mana woi... 😂😂😂😂	keputusan rakyat mana woi

Figure 7. Data Cleaning Results of TikTok

3.2.2 Case Folding

Case folding was applied to convert all characters in the text into lowercase format to standardize textual data across all samples. This process ensured consistency in word representation and minimized variations caused by differences in letter capitalization, which could otherwise lead to redundant or fragmented features during analysis. By unifying word forms regardless of their original case, case folding helped prevent unnecessary duplication of features and contributed to a more reliable and efficient feature extraction process, thereby supporting improved performance in subsequent sentiment classification tasks. The results of the case folding process applied to Instagram, X, and TikTok comments data are shown in Figures 8, 9, and 10.

1 to 10 of 10 entries Filter

index	text	case_folded
0	@dimasprasetyo18juli lo aja tanem diri dalam tanah	lo aja tanem diri dalam tanah
1		
2	buat orang2 yg ga pro ikn, percayalah suatu saat di masa depan, kalian bahkan anak cucu kalian bangga mempunyai ibukota seperti ikn🙌	buat orang2 yg ga pro ikn percayalah suatu saat di masa depan kalian bahkan anak cucu kalian bangga mempunyai ibukota seperti ikn
3	@deni_sup09 nandur o Dewe kebutuhan mu	nandur o dewe kebutuhan mu
4	Pak kalau butuh tukang AC atau tukang listrik di sana, kabar2 gih pak..🙏🙏😊	pak kalau butuh tukang ac atau tukang listrik di sana kabar2 gih pak
5	Semangat.terus.paak terimakasih kepada bapak jokowi telah.memajukan.indonesia.	semangatteruspaak terimakasih kepada bapak jokowi telahmemajukanindonesia
6	@adisk2912 iyow... Tak wa ne bapak e, sing tenang	iyow tak wa ne bapak e sing tenang
7	Terbaik pa	terbaik pa
8	Sehat selalu pak Jokowi..... JBU	sehat selalu pak jokowi jbu
9	Semoga di kota baru nanti gak ada yg buang sampah sembarangan	semoga di kota baru nanti gak ada yg buang sampah sembarangan

Figure 8. Case Folding Results of Instagram

Journal of Information System Research (JOSH)
Volume 7, No. 2, Januari 2026, pp 349−359
ISSN 2686-228X (media online)
https://ejurnal.seminar-id.com/index.php/josh/
DOI 10.47065/josh.v7i2.8839

1 to 10 of 10 entries Filter

index	full_text	case_folded
0	@O5Bravo Pemindahan Ibukota itu ok tapi kalau model kaya IKN. Iya lucu saja. Alasannya tiap waktu saja semakin konyol. Semakin tidak waras saja orang yang bela IKN. Saya voter 03. Tapi anggap mereka orang gila saja dukung proyek mangkrak. Sudahlah mangkrak. Biaya sudah habis banyak.	pemindahan ibukota itu ok tapi kalau model kaya ikn iya lucu saja alasannya tiap waktu saja semakin konyol semakin tidak waras saja orang yang bela ikn saya voter 03 tapi anggap mereka orang gila saja dukung proyek mangkrak sudahlah mangkrak biaya sudah habis banyak
1	Gubernur Kaltim : IKN akan Beroperasi Paling Lambat 2028 ... Pak @prabowo kenapa belum juga teken perpres pemindahan ASN ke IKN? https://t.co/Nn7dOWcnzv	gubernur kaltim ikn akan beroperasi paling lambat 2028 pak kenapa belum juga teken perpres pemindahan asn ke ikn
2	Makanya mata rantainya diputus dengan pemindahan ibukota. Gara gara itu Kelas liberal dan Islamis modern benci banget sama IKN. Padahal mereka ini minoritas. Di pilpres aja pemilih capres pro IKN jumlahnya 58+16 persen kok. https://t.co/1TwHrYCjGO	makanya mata rantainya diputus dengan pemindahan ibukota gara gara itu kelas liberal dan islamis modern benci banget sama ikn padahal mereka ini minoritas di pilpres aja pemilih capres pro ikn jumlahnya 5816 persen kok
3	@zhil_arf Barangkali itu salah satu motivasi pemindahan IKN ke Kaltim krn di sana kaum progresif liberal yg menamakan diri civil society maupun Islamis garis keras tidak terlalu kuat.	barangkali itu salah satu motivasi pemindahan ikn ke kaltim krn di sana kaum progresif liberal yg menamakan diri civil society maupun islamis garis keras tidak terlalu kuat
4	Waisak 2025 di Kaltim: Doa Damai untuk Dunia dan Kelancaran Proses Pemindahan IKN https://t.co/ERxWDON1Fp	waisak 2025 di kaltim doa damai untuk dunia dan kelancaran proses pemindahan ikn
5	@zhambernyowo @sarah_pndjtn Jadi??? Yg km 50 dia presiden nya. Yg janji pesan mobil dia. Yg ambisi pemindahan IKN dia. Yg mengatakan IPK kurang 2 dia. Pemuja nya keras keras ga menentu. Yg di puja lembek lembek bau. Zholim....	jadi yg km 50 dia presiden nya yg janji pesan mobil dia yg ambisi pemindahan ikn dia yg mengatakan ipk kurang 2 dia pemuja nya keras keras ga menentu yg di puja lembek lembek bau zholim
6	Pemindahan IKN merupakan wujud nyata pemerintah dalam meningkatkan kesejahteraan dan mempererat persatuan seluruh masyarakat Indonesia. https://t.co/YGPJ17d0sN	pemindahan ikn merupakan wujud nyata pemerintah dalam meningkatkan kesejahteraan dan mempererat persatuan seluruh masyarakat indonesia
7	Pemutarbalikan Kebijakan Populer Memutarbalikkan pencapaian seperti pemindahan IKN dan pembangunan infrastruktur menyebutnya sebagai pemborosan atau upaya pencitraan untuk mengurangi dukungan publik.	pemutarbalikan kebijakan populer memutarbalikkan pencapaian seperti pemindahan ikn dan pembangunan infrastruktur menyebutnya sebagai pemborosan atau upaya pencitraan untuk mengurangi dukungan publik
8	Opini saya lainnya terbit di Koran terpopuler di Jawa Tengah Suara Merdeka edisi 8/5/2025 terkait Karut Marut Pemindahan Aparatur Negara ke IKN. Dapat diakses melalui link berikut https://t.co/6oNB0sAQ23. #NicholasMartuaSiagian #Koran #SuaraMerdeka https://t.co/RJlWAzSiGR	opini saya lainnya terbit di koran terpopuler di jawa tengah suara merdeka edisi 852025 terkait karut marut pemindahan aparatur negara ke ikn dapat diakses melalui link berikut
9	Komisi II DPR RI : Kaji Kesiapan Infrastruktur Terkait Pemindahan ASN Ke IKN - TVR 120 https://t.co/rFCyMVTcAM	komisi ii dpr ri kaji kesiapan infrastruktur terkait pemindahan asn ke ikn tvr 120

Figure 9. Case Folding Results of X

1 to 10 of 10 entries Filter

index	full_text	case_folded
0		
1	❤️❤️❤️❤️❤️	❤️❤️❤️❤️❤️
2	😊	😊
3	😂😂😂😂😂	😂😂😂😂😂
4	😂😂😂	😂😂😂
5	hmmmmmmm	hmmmmmmm
6	cot cot	cot cot
7	enteng banget tu rahang.😡😡😡	enteng banget tu rahang.😡😡😡
8	proyek Mulyono bkn Jokowi	proyek mulyono bkn jokowi
9	keputusan rakyat mana woi... 😂😂😂😂	keputusan rakyat mana woi... 😂😂😂😂

Figure 10. Case Folding Results of TikTok

3.2.3 Tokenization

Tokenization divided the cleaned text into individual tokens or words by segmenting continuous text sequences into discrete and meaningful linguistic units. This process transformed unstructured textual data into a structured format that could be effectively processed by machine learning algorithms. By enabling the model to analyze text at the word level, tokenization made it possible to capture the frequency, distribution, and co-occurrence of terms across the dataset.

As a result, important sentiment-related expressions, keywords, and contextual patterns could be more clearly identified and analyzed, allowing the model to better capture variations in opinion and emotional tone across different comments. Furthermore, tokenization served as a crucial foundation for subsequent text representation and feature extraction techniques, such as term weighting and vectorization, which play an essential role in improving the performance and accuracy of sentiment classification models. The results of the tokenization process for comments data from platform X, TikTok, and Instagram are presented in Figures 11, 12, and 13.

	full_text	tokenized
0	@O5Bravo Pemindahan Ibukota itu ok tapi kalau ...	[pindah, ibukota, ok, model, kaya, ikn, iya, l...
1	Gubernur Kaltim : IKN akan Beroperasi Paling L...	[gubernur, kaltim, ikn, operasi, lambat, 2028,...
2	Makanya mata rantainya diputus dengan pemindah...	[mata, rantai, putus, pindah, ibukota, gara, g...
3	@zhil_arf Barangkali itu salah satu motivasi p...	[barangkali, salah, motivasi, pindah, ikn, kal...
4	Waisak 2025 di Kaltim: Doa Damai untuk Dunia d...	[waisak, 2025, kaltim, doa, damai, dunia, lanc...
5	@zhambernyowo @sarah_pndjtn Jadi??? Yg km 50 d...	[yg, km, 50, presiden, nya, yg, janji, pesan, ...
6	Pemindahan IKN merupakan wujud nyata pemerinta...	[pindah, ikn, wujud, nyata, perintah, tingkat,...
7	Pemutarbalikan Kebijakan Populer Memutarbalikk...	[pemutarbalikan, bijak, populer, memutarbalikk...
8	Opini saya lainnya terbit di Koran terpopuler ...	[opini, terbit, koran, populer, jawa, suara, m...
9	Komisi II DPR RI : Kaji Kesiapan Infrastruktur...	[komisi, ii, dpr, ri, kaji, kesiap, infrastruk...

Figure 11. Tokenization Results of X

Journal of Information System Research (JOSH)
Volume 7, No. 2, Januari 2026, pp 349−359
ISSN 2686-228X (media online)
https://ejurnal.seminar-id.com/index.php/josh/
DOI 10.47065/josh.v7i2.8839

	full_text	tokenized
0		[]
1	❤️❤️❤️❤️❤️	[]
2	😊	[]
3	😂😂😂😂😂	[]
4	😂😂😂	[]
5	hmmmmmmm	[hmmmmmmm]
6	cot cot	[cot, cot]
7	enteng banget tu rahang.🤩🤩🤩	[enteng, banget, tu, rahang]
8	proyek Mulyono bkn Jokowi	[proyek, mulyono, bkn, jokowi]
9	keputusan rakyat mana woi... 😂😂😂😂	[putus, rakyat, woi]

Figure 12. Tokenization Results of TikTok

	text	tokenized
0	@dimasprasetyo18juli lo aja tanem diri dalam t...	[lo, aja, tanem, tanah]
1		[]
2	buat orang2 yg ga pro ikn, percayalah suatu sa...	[orang2, yg, ga, pro, ikn, percaya, anak, cucu...
3	@deni_sup09 nandur o Dewe kebutuhan mu	[nandur, o, dewe, butuh, mu]
4	Pak kalau butuh tukang AC atau tukang listrik ...	[butuh, tukang, ac, tukang, listrik, kabar2, gih]
5	Semangat.terus.paak terimakasih kepada bapak j...	[semangatteruspaak, terimakasih, jokowi, telah...
6	@adisk2912 iyow... Tak wa ne bapak e, sing tenang	[iyow, wa, ne, e, sing, tenang]
7	Terbaik pa	[baik, pa]
8	Sehat selalu pak Jokowi..... JBU	[sehat, jokowi, jbu]
9	Semoga di kota baru nanti gak ada yg buang sam...	[moga, kota, gak, yg, buang, sampah, sembarang]

Figure 13. Tokenization Results of Instagram

3.2.4 Stopword Removal

Stopword removal was performed to eliminate commonly used words that carry little semantic value, such as conjunctions and prepositions. Removing stopwords reduced dimensionality and improved computational efficiency without significantly affecting sentiment-related information. Figures 14,15, and 16 shows the results of the stopword removal process for Instagram, X, and TikTok comments data.

1 to 10 of 10 entries Filter

index	text	stopword_removed
0	@dimasprasetyo18juli lo aja tanem diri dalam tanah	@dimasprasetyo18juli lo aja tanem tanah
1		
2	buat orang2 yg ga pro ikn, percayalah suatu saat di masa depan, kalian bahkan anak cucu kalian bangga mempunyai ibukota seperti ikn🙌	orang2 yg ga pro ikn, percayalah depan, anak cucu bangga ibukota ikn🙌
3	@deni_sup09 nandur o Dewe kebutuhan mu	@deni_sup09 nandur o dewe kebutuhan mu
4	Pak kalau butuh tukang AC atau tukang listrik di sana, kabar2 gih pak..🙏🙏😊	butuh tukang ac tukang listrik sana, kabar2 gih pak..🙏🙏😊
5	Semangat.terus.paak terimakasih kepada bapak jokowi telah.memajukan.indonesia.	semangat.terus.paak terimakasih jokowi telah.memajukan.indonesia.
6	@adisk2912 iyow... Tak wa ne bapak e, sing tenang	@adisk2912 iyow... wa ne e, sing tenang
7	Terbaik pa	terbaik pa
8	Sehat selalu pak Jokowi..... JBU	sehat jokowi..... jbu
9	Semoga di kota baru nanti gak ada yg buang sampah sembarangan	semoga kota gak yg buang sampah sembarangan

Figure 14. Stopword Removal Results of Instagram

1 to 10 of 10 entries Filter

index	full_text	stopword_removed
0	@O5Bravo Pemindahan Ibukota itu ok tapi kalau model kaya IKN. Iya lucu saja. Alasannya tiap waktu saja semakin konyol. Semakin tidak waras saja orang yang bela IKN. Saya voter 03. Tapi anggap mereka orang gila saja dukung proyek mangkrak. Sudahlah mangkrak. Biaya sudah habis banyak.	@o5bravo pemindahan ibukota ok model kaya ikn. iya lucu saja. alasannya konyol. waras orang bela ikn. voter 03. anggap orang gila dukung proyek mangkrak. mangkrak. biaya habis banyak.
1	Gubernur Kaltim : IKN akan Beroperasi Paling Lambat 2028 ... Pak @prabowo kenapa belum juga teken perpres pemindahan ASN ke IKN? https://t.co/Nn7dOWcnzv	gubernur kaltim : ikn beroperasi lambat 2028 ... @prabowo teken perpres pemindahan asn ikn? https://t.co/nn7dowcnzv
2	Makanya mata rantainya diputus dengan pemindahan ibukota. Gara gara itu Kelas liberal dan Islamis modern benci banget sama IKN. Padahal mereka ini minoritas. Di pilpres aja pemilih capres pro IKN jumlahnya 58+16 persen kok. https://t.co/1TwHrYCjGO	mata rantainya diputus pemindahan ibukota. gara gara kelas liberal islamis modern benci banget ikn. minoritas. pilpres aja pemilih capres pro ikn 58+16 persen kok. https://t.co/1twhrycjgo
3	@zhil_arf Barangkali itu salah satu motivasi pemindahan IKN ke Kaltim krn di sana kaum progresif liberal yg menamakan diri civil society maupun Islamis garis keras tidak terlalu kuat.	@zhil_arf barangkali salah motivasi pemindahan ikn kaltim krn kaum progresif liberal yg menamakan civil society islamis garis keras kuat.
4	Waisak 2025 di Kaltim: Doa Damai untuk Dunia dan Kelancaran Proses Pemindahan IKN https://t.co/ERxWDON1Fp	waisak 2025 kaltim: doa damai dunia kelancaran proses pemindahan ikn https://t.co/erxwdon1fp
5	@zhambernyowo @sarah_pndjtn Jadi??? Yg km 50 dia presiden nya. Yg janji pesan mobil dia. Yg ambisi pemindahan IKN dia. Yg mengatakan IPK kurang 2 dia. Pemuja nya keras keras ga menentu. Yg di puja lembek lembek bau. Zholim....	@zhambernyowo @sarah_pndjtn jadi??? yg km 50 presiden nya. yg janji pesan mobil dia. yg ambisi pemindahan ikn dia. yg ipk 2 dia. pemuja nya keras keras ga menentu. yg puja lembek lembek bau. zholim....
6	Pemindahan IKN merupakan wujud nyata pemerintah dalam meningkatkan kesejahteraan dan mempererat persatuan seluruh masyarakat Indonesia. https://t.co/YGPJ17d0sN	pemindahan ikn wujud nyata pemerintah meningkatkan kesejahteraan mempererat persatuan masyarakat indonesia. https://t.co/ygpj17d0sn
7	Pemutarbalikan Kebijakan Populer Memutarbalikkan pencapaian seperti pemindahan IKN dan pembangunan infrastruktur menyebutnya sebagai pemborosan atau upaya pencitraan untuk mengurangi dukungan publik.	pemutarbalikan kebijakan populer memutarbalikkan pencapaian pemindahan ikn pembangunan infrastruktur menyebutnya pemborosan upaya pencitraan mengurangi dukungan publik.
8	Opini saya lainnya terbit di Koran terpopuler di Jawa Tengah Suara Merdeka edisi 8/5/2025 terkait Karut Marut Pemindahan Aparatur Negara ke IKN. Dapat diakses melalui link berikut https://t.co/6oNB0sAQ23. #NicholasMartuaSiagian #Koran #SuaraMerdeka https://t.co/RJlWAzSiGR	opini terbit koran terpopuler jawa suara merdeka edisi 8/5/2025 terkait karut marut pemindahan aparatur negara ikn. diakses link https://t.co/6onb0saq23. #nicholasmartuasiagian #koran #suaramerdeka https://t.co/rjlwazsigr
9	Komisi II DPR RI : Kaji Kesiapan Infrastruktur Terkait Pemindahan ASN Ke IKN - TVR 120 https://t.co/rFCyMVTcAM	komisi ii dpr ri : kaji kesiapan infrastruktur terkait pemindahan asn ikn - tvr 120 https://t.co/rfcymvtcam

Figure 15. Stopword Removal Results of X

Journal of Information System Research (JOSH)
Volume 7, No. 2, Januari 2026, pp 349−359
ISSN 2686-228X (media online)
https://ejurnal.seminar-id.com/index.php/josh/
DOI 10.47065/josh.v7i2.8839

1 to 10 of 10 entries Filter

index	full_text	stopword_removed
0		
1	❤️❤️❤️❤️❤️	❤️❤️❤️❤️❤️
2	😊	😊
3	😂😂😂😂😂	😂😂😂😂😂
4	😂😂😂	😂😂😂
5	hmmmmmmm	hmmmmmmm
6	cot cot	cot cot
7	enteng banget tu rahang.🤣🤣🤣	enteng banget tu rahang.🤣🤣🤣
8	proyek Mulyono bkn Jokowi	proyek mulyono bkn jokowi
9	keputusan rakyat mana woi... 😂😂😂😂	keputusan rakyat woi... 😂😂😂😂

Figure 16. Stopword Removal Results of TikTok

3.2.5 Stemming

Stemming was applied using the Sastrawi algorithm to convert words into their root forms by removing affixes. This process reduced linguistic variations in Indonesian words and helped unify similar terms into a single representation, which improved classification performance. Figures 17, 18, and 19 presents the stemming results for Instagram, X, and TikTok comments data using the Sastrawi algorithm.

1 to 10 of 10 entries Filter

index	text	stemmed
0	@dimasprasetyo18juli lo aja tanem diri dalam tanah	dimasprasetyo18juli lo aja tanem tanah
1		
2	buat orang2 yg ga pro ikn, percayalah suatu saat di masa depan, kalian bahkan anak cucu kalian bangga mempunyai ibukota seperti ikn🙌	orang2 yg ga pro ikn percaya depan anak cucu bangga ibukota ikn
3	@deni_sup09 nandur o Dewe kebutuhan mu	den sup09 nandur o dewe butuh mu
4	Pak kalau butuh tukang AC atau tukang listrik di sana, kabar2 gih pak..🙏🙏😁	butuh tukang ac tukang listrik sana kabar2 gih pak
5	Semangat.terus.paak terimakasih kepada bapak jokowi telah.memajukan.indonesia.	semangat terus paak terimakasih jokowi telah maju indonesia
6	@adisk2912 iyow... Tak wa ne bapak e, sing tenang	adisk2912 iyow wa ne e sing tenang
7	Terbaik pa	baik pa
8	Sehat selalu pak Jokowi..... JBU	sehat jokowi jbu
9	Semoga di kota baru nanti gak ada yg buang sampah sembarangan	moga kota gak yg buang sampah sembarang

Figure 17. Stemming Results of Instagram

1 to 10 of 10 entries Filter

index	full_text	stemmed
0	@O5Bravo Pemindahan Ibukota itu ok tapi kalau model kaya IKN. Iya lucu saja. Alasannya tiap waktu saja semakin konyol. Semakin tidak waras saja orang yang bela IKN. Saya voter 03. Tapi anggap mereka orang gila saja dukung proyek mangkrak. Sudahlah mangkrak. Biaya sudah habis banyak.	o5bravo pindah ibukota ok model kaya ikn iya lucu saja alas konyol waras orang bela ikn voter 03 anggap orang gila dukung proyek mangkrak mangkrak biaya habis banyak
1	Gubernur Kaltim : IKN akan Beroperasi Paling Lambat 2028 ... Pak @prabowo kenapa belum juga teken perpres pemindahan ASN ke IKN? https://t.co/Nn7dOWcnzv	gubernur kaltim ikn operasi lambat 2028 prabowo teken pres pindah asn ikn https t co nn7dowcnzv
2	Makanya mata rantainya diputus dengan pemindahan ibukota. Gara gara itu Kelas liberal dan Islamis modern benci banget sama IKN. Padahal mereka ini minoritas. Di pilpres aja pemilih capres pro IKN jumlahnya 58+16 persen kok. https://t.co/1TwHrYCjGO	mata rantai putus pindah ibukota gara gara kelas liberal islamis modern benci banget ikn minoritas pilpres aja pilih capres pro ikn 58 16 persen kok https t co 1twhrycjgo
3	@zhil_arf Barangkali itu salah satu motivasi pemindahan IKN ke Kaltim krn di sana kaum progresif liberal yg menamakan diri civil society maupun Islamis garis keras tidak terlalu kuat.	zhil arf barangkali salah motivasi pindah ikn kaltim krn kaum progresif liberal yg nama civil society islamis garis keras kuat
4	Waisak 2025 di Kaltim: Doa Damai untuk Dunia dan Kelancaran Proses Pemindahan IKN https://t.co/ERxWDON1Fp	waisak 2025 kaltim doa damai dunia lancar proses pindah ikn https t co erxwdon1fp
5	@zhambernyowo @sarah_pndjtn Jadi??? Yg km 50 dia presiden nya. Yg janji pesan mobil dia. Yg ambisi pemindahan IKN dia. Yg mengatakan IPK kurang 2 dia. Pemuja nya keras keras ga menentu. Yg di puja lembek lembek bau. Zholim....	zhambernyowo sarah pndjtn jadi yg km 50 presiden nya yg janji pesan mobil dia yg ambisi pindah ikn dia yg ipk 2 dia puja nya keras keras ga tentu yg puja lembek lembek bau zholim
6	Pemindahan IKN merupakan wujud nyata pemerintah dalam meningkatkan kesejahteraan dan mempererat persatuan seluruh masyarakat Indonesia. https://t.co/YGPJ17d0sN	pindah ikn wujud nyata perintah tingkat sejahtera erat satu masyarakat indonesia https t co ygpj17d0sn
7	Pemutarbalikan Kebijakan Populer Memutarbalikkan pencapaian seperti pemindahan IKN dan pembangunan infrastruktur menyebutnya sebagai pemborosan atau upaya pencitraan untuk mengurangi dukungan publik.	pemutarbalikan bijak populer memutarbalikkan capai pindah ikn bangun infrastruktur sebut boros upaya citra kurang dukung publik
8	Opini saya lainnya terbit di Koran terpopuler di Jawa Tengah Suara Merdeka edisi 8/5/2025 terkait Karut Marut Pemindahan Aparatur Negara ke IKN. Dapat diakses melalui link berikut https://t.co/6oNB0sAQ23. #NicholasMartuaSiagian #Koran #SuaraMerdeka https://t.co/RJlWAzSiGR	opini terbit koran populer jawa suara merdeka edisi 8 5 2025 kait karut marut pindah aparatur negara ikn akses link https t co 6onb0saq23 nicholasmartuasiagian koran suaramerdeka https t co rjlwazsigr
9	Komisi II DPR RI : Kaji Kesiapan Infrastruktur Terkait Pemindahan ASN Ke IKN - TVR 120 https://t.co/rFCyMVTcAM	komisi ii dpr ri kaji kesiap infrastruktur kait pindah asn ikn - tvr 120 https t co rfcymvtcam

Figure 18. Stemming Results of X

1 to 10 of 10 entries Filter

index	full_text	stemmed
0		
1	❤️❤️❤️❤️❤️	
2	😊	
3	😂😂😂😂😂	
4	😂😂😂	
5	hmmmmmmm	hmmmmmmm
6	cot cot	cot cot
7	enteng banget tu rahang.🤣🤣🤣	enteng banget tu rahang
8	proyek Mulyono bkn Jokowi	proyek mulyono bkn jokowi
9	keputusan rakyat mana woi... 😂😂😂😂	putus rakyat woi

Figure 19. Stemming Results of TikTok

3.3 Feature Extraction and Classification Results

After preprocessing, the cleaned text data were transformed into numerical features using the Term Frequency–Inverse Document Frequency method. This technique assigns weights to words based on their frequency within individual documents and their distribution across the dataset, enabling the model to capture important terms

Journal of Information System Research (JOSH)
Volume 7, No. 2, Januari 2026, pp 349−359
ISSN 2686-228X (media online)
https://ejurnal.seminar-id.com/index.php/josh/
DOI 10.47065/josh.v7i2.8839

related to public sentiment on the capital relocation issue. The extracted TF-IDF features were then used as input for the sentiment classification process using the Support Vector Machine algorithm with a linear kernel. The dataset was divided into training and testing sets using an 80:20 split. Hyperparameter tuning was performed using grid search to determine the optimal regularization parameter, ensuring optimal model performance. The classification results demonstrate that the Support Vector Machine model is capable of effectively distinguishing sentiment categories across all three platforms, despite differences in language style and content characteristics.

3.4 Evaluation and Discussion of Results

The sentiment distribution for each social media platform based on the classification results is shown in Figure 20 and Figure 21 illustrates the comparison of sentiment distribution across X, TikTok, and Instagram using a bar chart.

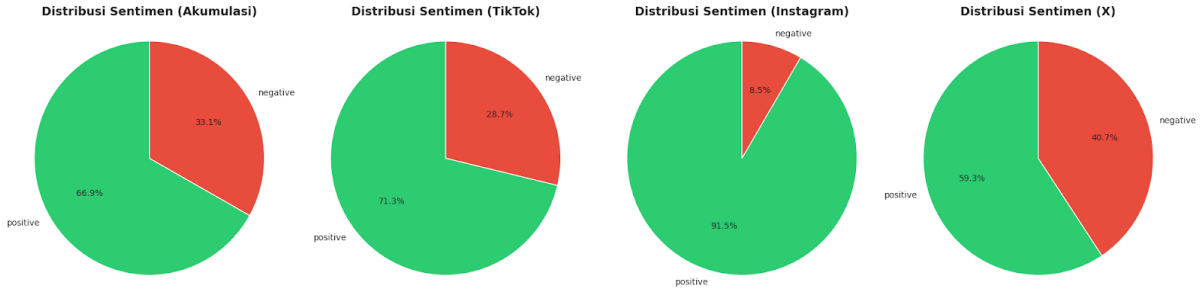

Figure 20. Distribution of sentiment for each platform using a pie chart.

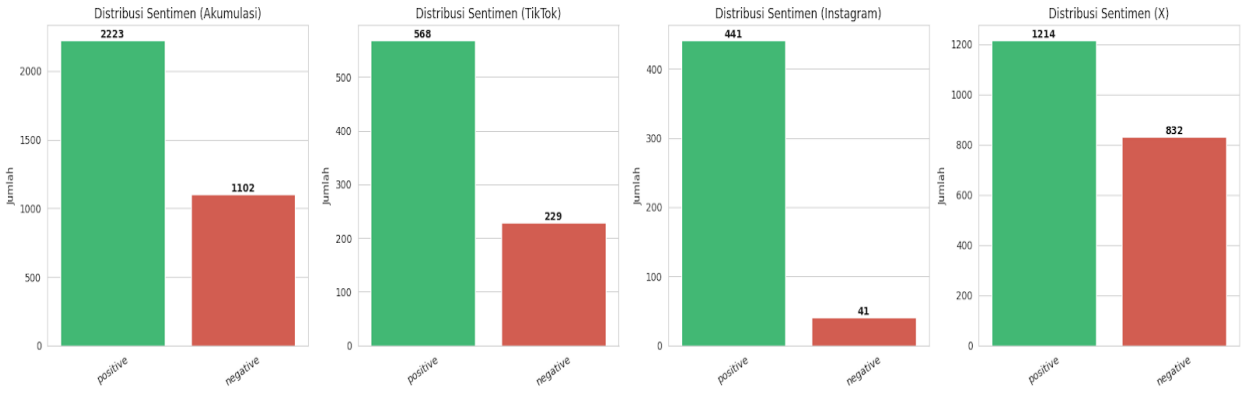

Figure 21. Image of sentiment distribution for each platform using a bar chart

3.4 Evaluation and Discussion of Results

Model evaluation shows that the Support Vector Machine achieved an accuracy of 90.23% with a macro F1-score of 0.8905, indicating strong classification performance. The accuracy obtained in this study is consistent with previous sentiment analysis research on the Indonesian capital relocation using Support Vector Machine, which also reported high classification performance on Twitter data [18]. The confusion matrix analysis reveals that the model performs better in identifying positive sentiment compared to negative sentiment, which is partly influenced by the imbalance in sentiment distribution within the dataset. Sentiment distribution analysis indicates platform-specific differences. Instagram and TikTok exhibit predominantly positive sentiment, reflecting the visual and emotional nature of these platforms, while X shows a higher proportion of negative sentiment, consistent with its role as a space for critical political discourse. These findings confirm that user demographics, platform culture, and content formats significantly influence how public opinions are expressed [20]. The results also demonstrate that the TF-IDF-based Support Vector Machine approach outperforms several traditional classification methods reported in previous studies, reinforcing its suitability for sentiment analysis related to public policy issues [21]. However, limitations such as data imbalance and the exclusion of comparative algorithms suggest the need for further research to improve robustness and generalizability. The evaluation results of the classification model using a confusion matrix are presented in Figure 22 and Figure 23 shows the classification report generated from the Support Vector Machine model used in this study.

Furthermore, the observed performance differences across sentiment classes highlight important methodological implications for future sentiment analysis studies. The relatively lower performance in identifying negative sentiment suggests that negative expressions may be more linguistically diverse, implicit, or context-dependent, making them more challenging for the model to classify accurately. This issue is commonly encountered in social media-based sentiment analysis, where sarcasm, irony, and ambiguous language frequently appear, particularly in political discussions. Addressing this challenge may require the incorporation of advanced

Journal of Information System Research (JOSH)
Volume 7, No. 2, Januari 2026, pp 349−359
ISSN 2686-228X (media online)
https://ejurnal.seminar-id.com/index.php/josh/
DOI 10.47065/josh.v7i2.8839

preprocessing techniques, sentiment lexicons, or class-balancing strategies to improve minority class representation. Additionally, integrating semantic-based feature representations or comparing the Support Vector Machine with other machine learning and deep learning approaches could provide further insights into model robustness and performance trade-offs. From a practical perspective, the findings of this study demonstrate that sentiment analysis can serve as an effective tool for policymakers to monitor public responses to major national policies in near real time, enabling more informed decision-making and public communication strategies. Overall, this study contributes empirical evidence supporting the effectiveness of TF-IDF-based Support Vector Machine models for analyzing public sentiment on policy-related issues across multiple social media platforms.

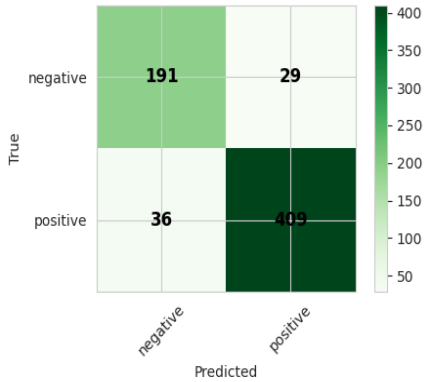

Figure 22. Confusion Matrix of SVM

```
=== HASIL UJI ===
Accuracy : 0.9023
F1-macro : 0.8905

Classification Report:
              precision    recall  f1-score   support

    negative       0.84      0.87      0.85       220
    positive       0.93      0.92      0.93       445

    accuracy                           0.90       665
   macro avg       0.89      0.89      0.89       665
weighted avg       0.90      0.90      0.90       665
```

Figure 23. Classification Report from SVM Model

4. CONCLUSION

This study concludes that the Support Vector Machine algorithm is effective and reliable for classifying public sentiment regarding the relocation of Indonesia's capital across X, TikTok, and Instagram. The results show that sentiment tends to be more positive on visually oriented platforms such as Instagram and TikTok, while X exhibits a higher proportion of critical opinions, reflecting differences in user characteristics and platform culture. Unlike earlier studies that focused on comparing multiple classifiers such as Naïve Bayes and K-Nearest Neighbor on platform X, this research emphasizes cross-platform sentiment analysis using Support Vector Machine as the primary classifier [19]. Text preprocessing and feature extraction using the Term FrequencyInverse Document Frequency method successfully supported sentiment classification, although data imbalance affected performance in minority sentiment classes. Despite achieving satisfactory accuracy, this study is limited by imbalanced data, a restricted collection period, and the use of a single classification method, indicating the need for future research with larger datasets, comparative algorithms, and broader policy topics.

REFERENCES

[1] "Asosiasi Penyelenggara Jasa Internet Indonesia." Accessed: Dec. 28, 2025. [Online]. Available: https://apjii.or.id/berita/d/apjii-jumlah-pengguna-internet-indonesia-tembus-221-juta-orang
[2] M. K. Saraswati and E. A. W. Adi, "Pemindahan Ibu Kota Negara Ke Provinsi Kalimantan Timur Berdasarkan Analisis SWOT," JISIP (Jurnal Ilmu Sos. dan Pendidikan), vol. 6, no. 2, pp. 4042–4052, 2022, doi: 10.58258/jisip.v6i2.3086.
[3] A. Yusuf, A. Rizani, R. Fitri, K. N. P. Pamungkas, and W. A. Saputra, "Sentimen Positif Atau Negatif: Perspektif Masyarakat Terhadap Pemindahan Ibu Kota Negara," J. Masy. Indones., vol. 50, no. 2, pp. 277–300, 2024, doi: 10.55981/jmi.2024.8842.
[4] S. Rihastuti, A. Rosyidi, A. Surakarta, and R. Forest, "PROGRES PEMBANGUNAN IKN DENGAN METODE RANDOM FOREST," J. Comput. Sci. Technol., vol. 5, no. 1, pp. 19–23, 2025, [Online]. Available: https://doi.org/10.54840/jcstech.v5i1.345
[5] N. Istamala, N. Azizah, O. Nurahim, and D. Daryono, "Opini Publik Berdasarkan Teori Agenda Setting Pada Proses Perencanaan Pemindahan IKN," J. Manaj. Sos. Ekon., vol. 4, no. 2, pp. 74–87, 2024, [Online]. Available:

Journal of Information System Research (JOSH)
Volume 7, No. 2, Januari 2026, pp 349−359
ISSN 2686-228X (media online)
https://ejurnal.seminar-id.com/index.php/josh/
DOI 10.47065/josh.v7i2.8839

https://doi.org/10.51903/dinamika.v4i2.517

[6] A. M. Siregar, "Analisis Sentimen Pindah Ibu Kota Negara (IKN) Baru pada Twitter Menggunakan Algoritma Naive Bayes dan Support Vector Machine (SVM)," Fakt. Exacta, vol. 16, no. 3, pp. 170–181, 2023, doi: 10.30998/faktorexacta.v16i3.16703.

[7] K. T. Amazio, "PERBANDINGAN NAÏVE BAYES CLASSIFIER DAN SUPPORT VECTOR MACHINE PADA ANALISIS SENTIMEN NETIZEN X # KABURAJADULU COMPARISON BETWEEN NAÏVE BAYES CLASSIFIER AND SUPPORT," Semin. Nas. Mhs. Fak. Teknol. Inf., vol. 4, no. September, pp. 389–397, 2025, [Online]. Available: https://journal.stiestekom.ac.id/index.php/dinamika/article/view/517

[8] I. Kamindang, M. Amijaya, and F. Ilmu Sosial dan Politik, "Tiktok Sebagai Media Komunikasi Politik Aktor Partai Politik Di Kota Palu," J. Ilmu Komun. UHO J. Penelit. Kaji. Ilmu Sos. dan Inf., vol. 9, no. 1, pp. 1–15, 2024, [Online]. Available: http://jurnalilmukomunikasi.uho.ac.id/index.php/journal/indexDOI:http://dx.doi.org/10.52423/jikuho.v9i1.151

[9] G. A. Saputri and D. Alita, "Analisis Sentimen Twitter Terhadap Pemindahan Ibu Kota Negara Menggunakan Support Vector Machine," J. Inform. J. Pengemb. IT, vol. 9, no. 3, pp. 213–223, 2024, doi: 10.30591/jpit.v9i3.6612.

[10] A. Suharman, M. K. Sulaeman, T. Industri, U. Muhammadiyah, and P. Hamka, "Analisis Sentimen Pengguna Aplikasi Livin ' by Mandiri Menggunakan Metode Support Vector Machine (SVM) dengan Ekstraksi Fitur TF-IDF dan Word2Vec User Sentiment Analysis of the Livin ' by Mandiri Application Using the Support Vector Machine (SVM)" JPTI (Jurnal Pendidikan dan Teknologi Indonesia) vol. 5, no. 8, pp. 2201–2212, 2025, [Online]. Available: https://doi.org/10.52436/1.jpti.941

[11] S. Andini, R. Kurniawan, S. Anwar, and K. Cirebon, "Analisis Sentimen Pengguna X Mengenai Opini," JITET (Jurnal Inform. dan Tek. Elektro Ter., vol. 13, no. 2, pp. 665–671, 2025, [Online]. Available: https://journal.eng.unila.ac.id/index.php/jitet/article/view/6299

[12] D. Pateman, T. F. Prasetyo, and H. Sujadi, "Sentiment Analysis of Government on Tiktok and X Platforms With Svm and Smote Approach," JITK (Jurnal Ilmu Pengetah. dan Teknol. Komputer), vol. 10, no. 4, pp. 900–908, 2025, doi: 10.33480/jitk.v10i4.6645.

[13] B. Setiawan, "A Review of Sentiment Analysis Applications in Indonesia Between 2023-2024," J. Inf. Eng. Educ. Technol., vol. 8, no. 2, pp. 71–83, 2025, doi: 10.26740/jieet.v8n2.p71-83.

[14] D. Haliza and M. Ikhsan, "Sentiment Analysis on Public Perception of the Nusantara Capital on Social Media X Using Support Vector Machine (SVM) and K-Nearest Neighbor (K-NN) Methods," J. Appl. Informatics Comput., vol. 9, no. 3, pp. 716–723, 2025, doi: 10.30871/jaic.v9i3.9318.

[15] E. W. P. Hanifah Afkar Nabila, "PERBANDINGAN ALGORITMA MACHINE LEARNING: SVM, RANDOM FOREST, DAN XGBOOST UNTUK PREDIKSI STROKE," RABIT J. Teknol. dan Sist. Inf. Univrab, vol. 10, no. 2, pp. 1098–1110, 2025, [Online]. Available: https://jurnal.univrab.ac.id/index.php/rabit/article/view/6444

[16] N. Hadi and D. Sugiarto, "Analisis Sentimen Pembangunan IKN pada Media Sosial X Menggunakan Algoritma SVM, Logistic Regression dan Naïve Bayes," J. Inform. J. Pengemb. IT, vol. 10, no. 1, pp. 37–49, 2025, doi: 10.30591/jpit.v10i1.7106.

[17] A. Setiawan and R. R. Suryono, "Analisis Sentimen Ibu Kota Nusantara menggunakan Algoritma Support Vector Machine dan Naïve Bayes," Edumatic J. Pendidik. Inform., vol. 8, no. 1, pp. 183–192, 2024, doi: 10.29408/edumatic.v8i1.25667.

[18] 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, 2025, doi: 10.33395/sinkron.v9i2.14622.

[19] A. Cahya, S. Zamani, Y. Nuryamin, and A. Priyatna, "Analisis Sentimen Pengguna TikTok tentang Pembangunan IKN Menggunakan Algoritma Naive Bayes dan Decision Tree," J. Nas. Teknol. Komput., vol. 5, pp. 1112–1123, 2025, [Online]. Available: https://publikasi.hawari.id/index.php/jnastek/article/view/323

[20] M. D. R. C. Priyanto, Azahari, and M. I. Sa'ad, "Analisis Sentimen Terhadap Kontroversi Pembangunan IKN di Media Sosial Twitter Menggunakan Metode Naïve Bayes," BIT (Bulletin Inf. Technol., vol. 6, no. 2, pp. 97–108, 2025, doi: 10.47065/bit.v5i2.1993.

[21] C. Huda and M. Betty Yel, "Analisa Sentimen Tentang Ibu Kota Nusantara (IKN) Dengan Menggunakan Algoritma K-Nearest Neighbors (KNN) dan Naïve Bayes," J. Ilmu Komput. dan Sist. Inf., vol. 7, no. 1, pp. 126–130, 2024, doi: 10.55338/jikomsi.v7i1.2846.