Proceeding – ICAMEB
International Conference on Accounting, Management, Entrepreneurship and Business (ICAMEB)

Sentiment Analysis of YouTube Comments on Indonesia's Capital Relocation Using Naive Bayes

1st Asro
Faculty of Information Technology,
STMIK IM
Bandung, Indonesia
asroharun6@gmail.com

2nd Novi Rukhviyanti
Faculty of Information Technology,
STMIK IM
Bandung, Indonesia
novi.rukhviyanti@stmik-im.ac.id

3rd Didik Santoso
Master enginering, Faculty of Science and Technology
Raharja University
Tangerang, Banten, Indonesia
didik.santoso@raharja.info

***Abstract*—**

The relocation of Indonesia's capital city from Jakarta to East Kalimantan has generated wide public debate, with YouTube becoming one of the main platforms where people actively express their opinions. This study explores those responses by applying the Naive Bayes algorithm to classify sentiments found in user comments. The comments were collected using Google's API and then refined through a series of text-cleaning steps, such as removing unnecessary words, applying normalization, and stemming. Once prepared, a Multinomial Naive Bayes classifier was employed to group the comments into three categories: positive, negative, and neutral. The results indicate stable performance, with the model correctly classifying about three out of four comments across various test splits. Negative sentiments were consistently identified with high reliability, while neutral and positive tones were more difficult to capture, as reflected in lower recall values. These findings highlight both the potential and the limitations of Naive Bayes in sentiment analysis, while also providing valuable insights into public opinion regarding Indonesia's capital relocation policy.

Keywords: *Sentiment Analysis, Naive Bayes, YouTube, Public Opinion, Capital Relocation.*

I. INTRODUCTION

The decision to relocate a nation's capital is a monumental policy that influences infrastructure, economic development, and social dynamics. Indonesia's plan to move its capital from Jakarta to East Kalimantan presents unique challenges, given that Jakarta has been the administrative center since the enactment of Law No. 10 of 1964

under President Soekarno (Cahya Herdiyani & Zailani, 2022). This shift has sparked extensive discussion across various media platforms, with YouTube emerging as a prominent space where citizens express support, criticism, and diverse perspectives regarding the policy. The sentiments voiced in YouTube comments can provide valuable insights for policymakers seeking to understand public responses to this transformative national initiative. This research investigates such sentiments using the Naive Bayes algorithm, a probability-based classification method widely recognized for its efficiency in handling large text datasets. The Multinomial Naive Bayes approach, grounded in Bayes' theorem, is often applied in Natural Language Processing (NLP) tasks (Alfiyanti, 2024). Comments were collected via the Google Client API and processed using Visual Studio, including text normalization, stopword removal, and stemming before classification. While the method demonstrated strong performance in recognizing negative comments, it showed limitations in capturing neutral and positive expressions accurately. The classification accuracy reached 75.64% with a 10% test split, 75.03% with a 20% split, and 74.69% with a 30% split, revealing the challenges of detecting nuanced sentiment variations.

The study underscores the significance of sentiment analysis not only in the context of public policy but also as a broader digital tool for monitoring real-time societal attitudes. In particular, debates on capital relocation resurfaced in 2019 when the issue trended on social media, and again on May 31, 2023, following President Joko Widodo's announcement of the official logo for the new capital, Ibu Kota Nusantara (IKN) (Putnarubun & Palembang, n.d.). These events generated polarized reactions, emphasizing the importance of sentiment mining for understanding public opinion on such impactful policies. Although Naive Bayes remains constrained by its probabilistic assumptions and limited ability to capture linguistic complexity, it continues to provide a solid foundation for text-based sentiment classification (Kinerja Naïve Bayes et al., n.d.). Historically, discussions about relocating Indonesia's capital have persisted since the administrations of Soekarno through Susilo Bambang Yudhoyono (Dhery et al., 2023). Beyond economic and infrastructural considerations, the relocation also carries cultural and social implications, potentially reshaping local identity and interactions between residents and newcomers. Consequently, exploring online sentiments offers a critical perspective on how communities perceive these transformations.

This research further highlights the role of modern information technology in large-scale data collection and analysis. Sentiment data can be retrieved through automated tools and APIs or compiled manually from digital sources (Mantik et al., 2023). By employing the Google Client API and Visual Studio, this study demonstrates how technological resources can facilitate complex sentiment analysis in the policy domain. Ultimately, the findings aim to contribute to the literature on digital public opinion analysis while providing practical insights for policymakers. The outcomes are expected to guide the development of more responsive communication strategies and support evidence-based decision-making regarding Indonesia's capital relocation.

Novelty and Rationale. Unlike most previous studies in Indonesia that focus on short texts from Twitter or Facebook, this research uniquely analyzes longer and context-rich comments from YouTube, which contain more argumentative and reflective expressions. This choice provides two key contributions: (i) it captures deeper public reasoning and emotional responses regarding the capital relocation policy, and (ii) it introduces a relatively rare primary data source in the Indonesian context of sentiment analysis. Practically, YouTube threads allow focused data collection from specific videos related to IKN, reducing off-topic noise while maintaining high relevance. Methodologically, the richer linguistic features in YouTube comments provide an ideal ground to test the Naive Bayes model's ability to handle long, informal, and multi-sentential structures. This distinguishing aspect enhances the study's originality and policy relevance by offering a new perspective on public opinion through discussion-level sentiment mining.

Previous studies on sentiment analysis concerning Indonesia's capital relocation have primarily focused on **Twitter** data and other short-text social media platforms. These works applied various machine learning algorithms such as Naive Bayes, Random Forest, Support Vector Machine (SVM), and more recent deep-learning architectures like Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM). While these approaches have contributed valuable insights into public opinion, most of them analyze short and fragmented messages, which often lack contextual depth and grammatical completeness. Consequently, they capture polarity (positive or negative) efficiently but fail to interpret complex linguistic expressions that appear in longer discussions. In addition, only a few recent studies have explored YouTube as a primary data source, even though it provides a richer environment for argument-based discourse and more comprehensive public reactions. Therefore, this research seeks to fill that gap by applying the Naive Bayes classifier on YouTube comments, comparing model performance under different test-train splits and discussing how long-form user opinions shape digital public sentiment about Indonesia's capital relocation policy.

To contextualize the development of this research area, **Table 1** summarizes several representative studies conducted between **2020 and 2025**, highlighting their datasets, applied algorithms, main findings, and the distinct contribution of the present study (Monica et al., 2024).

TABLE I. PREVIOUS RESEARCH ON SENTIMENT ANALYSIS

No	Study Title & Author(s)	Method / Platform	Main Findings & Relevance
1	*Twitter Sentiment Analysis of the Relocation of Indonesia's Capital City* (Sutoyo & Almaarif, 2020)	Twitter dataset; Naive Bayes, Logistic Regression, SVM, K-Nearest Neighbor	Identified a dominance of negative sentiments toward the capital relocation policy and provided an early overview of public perception using short Twitter texts.
2	*Sentiment Analysis of State Capital Relocation of Indonesia Using Convolutional Neural Network* (Welly Andi et al., 2025)	Twitter data; Deep Learning (CNN)	Applied a deep-learning approach to analyze public opinion on the relocation issue, showing methodological diversity beyond traditional machine-learning models.
3	*Sentiment Analysis on the Relocation of the National Capital (IKN) on Social Media X Using Naive Bayes and K-Nearest Neighbor* (Sutoyo & Almaarif, 2020)	Platform X (Twitter); Naive Bayes vs K-NN	Reported that Naive Bayes outperformed K-NN in classifying sentiments related to IKN discussions, confirming the efficiency of probabilistic models for text classification.
4	*Sentiment Analysis of Indonesia's Capital Relocation Using Word2Vec and Long Short-Term Memory (LSTM) Method* (Yanti & Utami, 2024)	Social-media data (Twitter and YouTube); Word2Vec + LSTM	Demonstrated a deep-learning approach that achieved strong accuracy for sentiment classification, highlighting advances in neural-network-based modeling for the IKN topic.
5	*Public Perception of IKN as the Capital City of Indonesia* (Rizki Fitrianto & Rizka Sugiarto, 2026)	Public-perception analysis using surveys and social-media data	Examined public attitudes toward the capital relocation from sociological and policy perspectives, offering broader contextual understanding beyond algorithmic analysis.

II. METHOD

This study adopts a **quantitative research approach** (Robi Padri, 2023),(A. N. S. Asro, 2024) to analyze public sentiment expressed in YouTube comments regarding Indonesia's capital city relocation. The **Naive Bayes** classification algorithm was chosen due to its proven efficiency in processing large-scale textual data and its ability to perform sentiment categorization with relatively low computational cost.

2.1 Operationalization of Variables

The main variable in this study is public sentiment toward Indonesia's capital relocation policy (IKN).
It is operationalized into three sentiment categories:

1. Positive – comments expressing support, optimism, or trust in the relocation policy.
2. Neutral – comments providing factual or descriptive statements without emotional tone.
3. Negative – comments containing criticism, disagreement, or concern regarding the policy.

These categories were determined using automatic classification by the Naive Bayes algorithm after text preprocessing.
Model performance was evaluated using accuracy, precision, recall, and F1-score, reflecting the reliability of the classification process.

2.2 Data Collection and Sampling Process

Data were obtained by crawling YouTube comments through the Google Client API. The collection period spanned several months, covering videos published from the official announcement of the capital relocation to the most recent discussions available at the time of research. A purposive sampling strategy was applied, selecting only videos that explicitly addressed the IKN relocation policy or related national news coverage (Zamzami et al., 2024),(H. L. A. Asro et al., n.d.),(H. L. A. Asro et al., 2024), Comments containing advertisements, spam, or irrelevant content were removed manually to maintain data quality. Each valid comment was collected along with metadata such as timestamp, user ID, and number of likes, supporting further relevance and engagement analysis.

2.3 Software and Hardware Requirements Analysis

A requirements analysis was carried out to ensure the system's efficiency and accuracy during implementation (Robi Padri et al., 2023),(Henderi et al., 2025). Software Requirements The project utilized several tools and libraries, as summarized in **Table 2**.

TABLE II. SOFTWARE REQUIREMENTS

No	Software	Function	Version
1	Google Client API	Collecting YouTube comment data	Latest
2	Visual Studio	Development environment	1.90
3	Python	Primary programming language	3.12
4	NLTK	Text preprocessing library	Latest
5	Scikit-learn	Naive Bayes implementation framework	Latest

Hardware Requirements The hardware configuration used in this research is shown in Table 2 (Fikry, 2023):

TABLE III. HARDWARE REQUIREMENTS

No	Hardware	Specification
1	Laptop	Lenovo
2	Processor	Intel Core i7
3	RAM	24 GB
4	Storage	512 GB SSD (high-speed I/O)

Each software component served a specific function: Google Client API for data extraction, NLTK for text preprocessing, and Scikit-learn for training and evaluating the Naive Bayes model.

2.4 Topic Selection and Data Preprocessing

The research focused specifically on public opinion related to Indonesia's capital relocation. After identifying relevant videos, a manual annotation process was conducted to exclude unrelated comments and retain only those directly discussing the policy. Valid comments were then categorized into positive, negative, or neutral sentiments. Data preprocessing was a critical phase to ensure textual consistency and analytical accuracy (Benedict & Rukhviyanti, 2025).

The main preprocessing steps included:

- Noise filtering: removing unnecessary elements such as URLs, emojis, and extraneous symbols.
- Tokenization: breaking sentences into individual tokens or lexical units.
- Stopword elimination: discarding frequently occurring words that carry minimal semantic value.
- Stemming and Lemmatization: converting words into their fundamental or root form to ensure consistency in text representation.
- These procedures enhance textual consistency and improve the effectiveness of sentiment classification.

2.5 Data Analysis and Evaluation Techniques

To evaluate the robustness of the Naive Bayes classifier, the dataset was divided into training and testing subsets using three ratios: 90% training / 10% testing, 80% training / 20% testing, and 70% training / 30% testing.

These ratios were selected to compare model consistency across different proportions of training data. A 90:10 split prioritizes model generalization with abundant training data, an 80:20 split serves as the standard reference in text classification tasks, and a 70:30 split tests model resilience under larger testing conditions.

The Multinomial Naive Bayes algorithm was implemented in Python using the Scikit-learn library, with features extracted through TF-IDF vectorization. Each sentiment class (positive, neutral, and negative) was trained and validated independently. Model evaluation employed standard classification metrics — accuracy, precision, recall, and F1-score — supported by a confusion matrix to visualize classification performance and identify potential misclassifications among sentiment categories. The evaluation results provided quantitative evidence of the model's capability to identify sentiment polarity across diverse textual expressions in YouTube comments.

2.6 Research Flow

The overall research process followed the sequential steps illustrated in **Figure 1**, beginning with topic identification and data collection, followed by preprocessing, model training, evaluation, and interpretation of results.

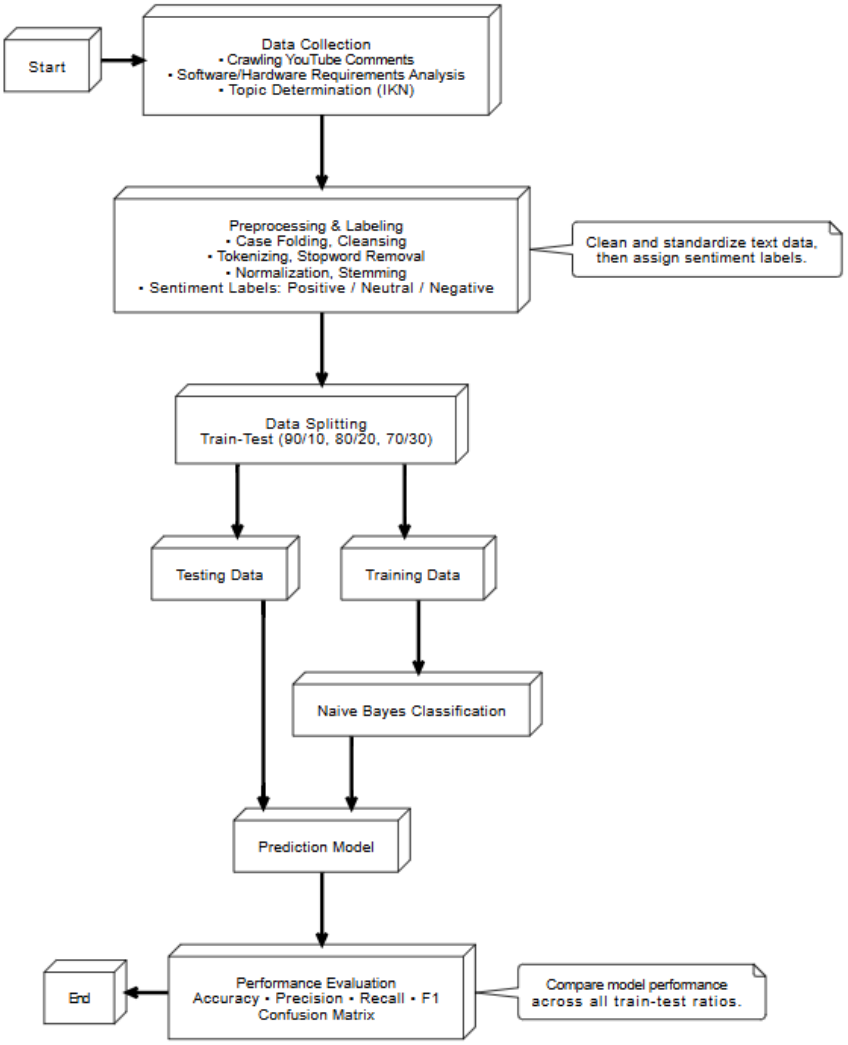


Figure 1. Research Methodology Flowchart

III. RESULTS AND DISCUSSION

This section presents the outcomes of sentiment analysis on YouTube comments related to the relocation of Indonesia's capital city (IKN) (Siregar, 2023), The study successfully collected and classified comments using the Naive Bayes approach, enabling deeper insights into public perceptions of the relocation policy.

Data Collection YouTube comments were gathered via the Google Client API over several months, beginning from the official announcement of the relocation until the latest phase of data gathering. The dataset was structured into tabular form for easier processing (Setiawan & Isnain, 2024).

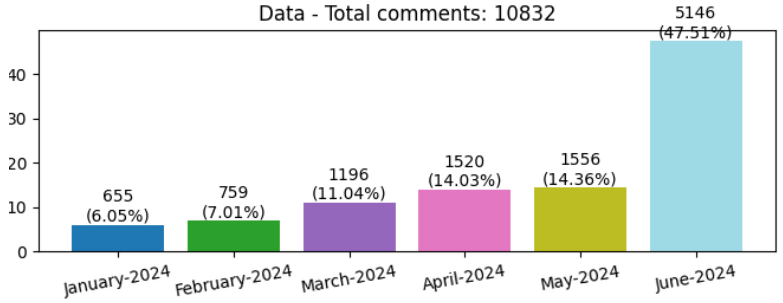


Figure 2 Total Data Collection

Figure 2 illustrates the monthly distribution of the 10,832 comments collected. The majority originated in June 2024 (47.51%), followed by May (14.36%) and April (14.03%). This trend suggests that public attention toward the relocation issue increased significantly over time.

Preprocessing Before analysis, the dataset underwent extensive preprocessing, including case folding, cleansing, tokenization, handling of non-standard words, stopword removal, and stemming (Jannah & Kusnawi, 2024),(Wibawa et al., 2019). These steps improved the consistency and quality of the textual data, which is crucial for sentiment classification. This the result of Sentiment Labeling and Distribution showing in figure 3

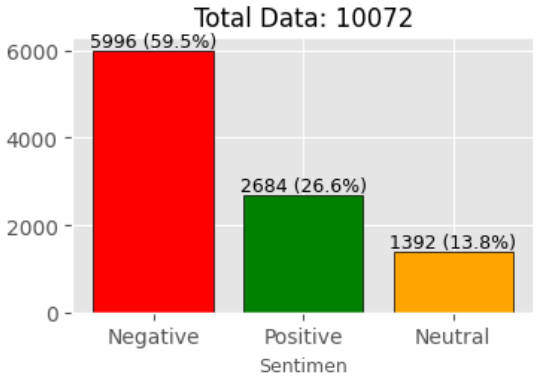


Figure 3. Distribution Sentiment

Manual labeling was performed to categorize comments as *positive*, *negative*, or *neutral*. A sentiment lexicon in Bahasa Indonesia was used as a reference. The distribution is summarized in Figure 3 **showing Negative:** 5,996 comments (59.5%), **Positive:** 2,684 comments (26.6%)**, Neutral:** 1,392 comments (13.8%) The predominance of negative sentiment highlights widespread skepticism and criticism surrounding the capital relocation policy.

Word Frequency Analysis

Figure 4. Word Clouds

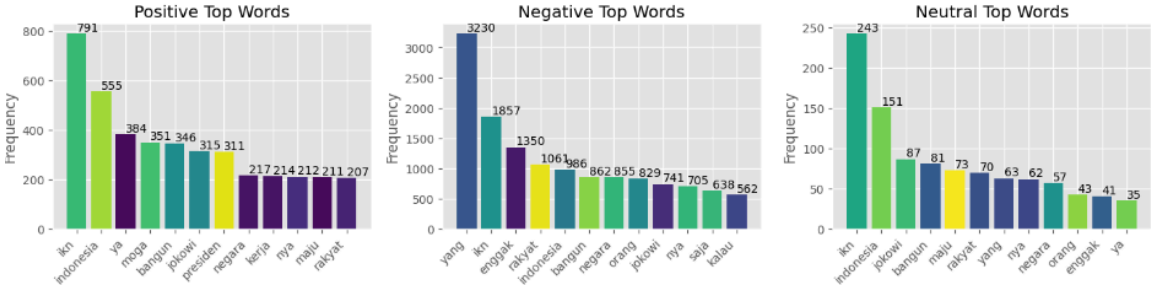


Figure 5. Frequency Top Words

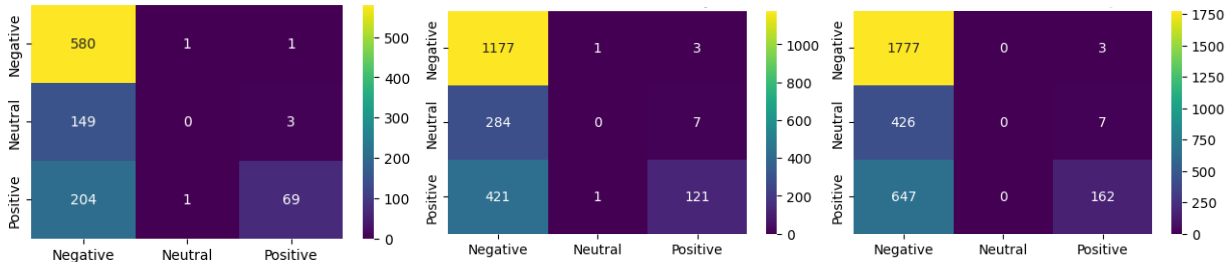
Word clouds (Figure 4) and bar charts (Figure 5) were generated to visualize the most frequent words within each sentiment category. Positive comments were dominated by terms like *"maju"* (progress), *"bangun"* (build), and *"Jokowi"*, reflecting optimism and support. Negative comments frequently contained *"yang"* (which), *"enggak"* (not), and *"rakyat"* (people), indicating dissatisfaction and disapproval. Neutral comments often used *"Indonesia"* and *"bangun"*, suggesting a more objective or balanced discussion. Quantitatively, the word *"yang"* appeared 3,230 times in negative comments, while *"Indonesia"* appeared 791 times in positive and 243 times in neutral comments. These patterns underscore the polarized nature of public discourse.

Naive Bayes Classification sentiment classification was performed using the Multinomial Naive Bayes algorithm (Fadlila Nurwanda et al., 2023). The probabilistic formula is given as:

$$P(C|X) = \frac{P(C).(C)}{P(X)}$$

where C represents sentiment class (positive, negative, or neutral) and X represents textual features extracted from comments.

Classification Results The evaluation was carried out with three different testing ratios (10%, 20%, 30%). summarizes precision, recall, F1-score, and accuracy. 10% test split: Accuracy = 75.64%; high precision for positive comments (0.945) but recall was very low (0.252). 20% test split: Accuracy = 75.03%; precision remained high (0.924), while recall dropped to 0.223. 30% test split: Accuracy = 74.69%; precision stayed at 0.942, but recall decreased further to 0.200. Across all splits, negative comments were consistently identified with high recall, while neutral comments were not detected at all, and positive comments showed strong precision but weak recall.



Figures 6 present the confusion matrices for each test split.

At 10% split, nearly all negative comments were correctly identified (580/582), but no neutral comments were detected, and only 25% of positive comments were classified correctly. At 20% split, the model correctly classified 1177/1181 negative comments, yet failed entirely on neutral comments and only captured 22% of positive comments. At 30% split, performance remained strong for negatives (1777/1780) but neutral detection failed again, with recall for positives falling to 20%. This shows a consistent weakness of Naive Bayes in distinguishing neutral and positive sentiments, even as dataset size increases.

Manual Performance Verification Manual calculations of accuracy, precision, recall, and F1-score confirmed the automated results: 10% split: Accuracy = 75.64%, Precision = 0.945, Recall = 0.252, F1 = 0.398, 20% split: Accuracy = 75.03%, Precision = 0.924, Recall = 0.223, F1 = 0.359, 30% split: Accuracy = 74.69%, Precision = 0.942, Recall = 0.200, F1 = 0.330.

The persistent imbalance between precision and recall demonstrates the algorithm's limitations in capturing sentiment diversity. Summary of Findings Overall, the evaluation demonstrates that the Naive Bayes classifier performs consistently in identifying negative sentiments with very high recall across all test splits (10%, 20%, and 30%). However, the model struggles to detect neutral sentiments, with zero correct classifications, and only partially identifies positive sentiments, where high precision is offset by very low recall. Manual calculations of accuracy, precision, recall, and F1-scores confirmed these limitations: accuracy remained around 75% in all scenarios, while F1-scores for positive comments fell below 0.40. These results highlight both the strengths and weaknesses of Naive Bayes—robust for identifying negative comments, but insufficiently sensitive for balanced sentiment coverage.

Summary of Findings: Overall, the evaluation demonstrates that the Naive Bayes classifier performs consistently in identifying negative sentiments with very high recall across all test splits (10%, 20%, and 30%). However, the model struggles to detect neutral sentiments, with zero correct classifications, and only partially identifies positive sentiments, where high precision is offset by very low recall. Manual calculations of accuracy, precision, recall, and F1-scores confirmed these limitations: accuracy remained around 75% in all scenarios, while F1-scores for positive comments fell below 0.40. These results highlight both the strengths and weaknesses of Naive Bayes—robust for identifying negative comments, but insufficiently sensitive for balanced sentiment coverage.

IV. CONCLUSIONS

This study successfully applied the Naive Bayes algorithm to analyze YouTube comments concerning the relocation of Indonesia's capital city from Jakarta to East Kalimantan. The results show that the model achieved consistently moderate to high performance in identifying public sentiments, particularly in detecting negative opinions related to the policy. However, the classifier exhibited noticeable limitations in recognizing positive and neutral expressions—while precision for positive sentiments was relatively strong, recall remained low, and neutral tones were often underrepresented. These findings highlight both the strengths and constraints of Naive Bayes in large-scale sentiment analysis. Although the model effectively captures critical perspectives and recurring negative tones, it struggles to interpret more nuanced or context-dependent language. Future research is encouraged to integrate Naive Bayes with complementary machine learning or deep learning techniques to enhance sentiment detection across more subtle emotional ranges, such as neutrality, irony, or sarcasm. From a practical perspective, the predominance of negative sentiment indicates considerable public skepticism toward the capital relocation project. Such insights are valuable for policymakers in refining public communication strategies, addressing public concerns, and strengthening positive narratives surrounding national development initiatives.

REFERENCES

Alfiyanti, D. (2024). Penerapan Algoritma Multinomial Naïve Bayes dan Support Vector Machine (SVM) untuk Analisis Sentimen terhadap Ulasan Google Maps di Taman Mini Indonesia Indah (TMII). *Jurnal SIMETRIS*, *15*(1).

Asro, A. N. S. (2024). Public Sentiment Analysis on the 2024 Presidential Election Using Naive Bayes Classifier (NBC) and Support Vector Machine (SVM) On Social Media Data. *PROSIDING SEMINAR INTERNASIONAL*. https://prosiding.amalinsani.org/

Asro, H. L. A., Wicaksono, A., & Herwanto, P. (n.d.). *Strategi Dinamis Menggunakan MongoDB untuk Analisis Sentimen terhadap Komentar YouTube Pilkada Gubernur Indonesia 2024 Dynamic Strategy Using MongoDB for Sentiment Analysis of YouTube Comments on the 2024 Indonesian Gubernatorial Elections*. https://jurnaldrpm.budiluhur.ac.id/index.php/Kresna/

Asro, H. L. A., Wicaksono, A., & Herwanto, P. (2024). *Strategi Dinamis Menggunakan MongoDB untuk Analisis Sentimen terhadap Komentar YouTube Pilkada Gubernur Indonesia 2024*. https://jurnaldrpm.budiluhur.ac.id/index.php/Kresna/

Benedict, K. V., & Rukhviyanti, N. (2025). Analysis of the Classification of Data on the Launch of Apple Mobile Phone Prices in China and Pakistan Using the Decision Tree Algorithm in Python Programming. *Eduvest-Journal of Universal Studies*, *5*(9). http://eduvest.greenvest.co.id

Cahya Herdiyani, T., & Zailani, A. U. (2022). Sentiment Analysis Terkait Pemindahan Ibu Kota Indonesia Menggunakan Metode Random Forest Berdasarkan Tweet Warga Negara Indonesia Sentiment Analysis Related to Transportation of Indonesian Capital City Using Random Forest Method Based On Tweet Of Indonesian Citizens. In *JTSI* (Vol. 3, Issue 2).

Dhery, H., Assyam, A., & Hasan, F. N. (2023). Analisis Sentimen Twitter Terhadap Perpindahan Ibu Kota Negara Ke IKN Nusantara Menggunakan Orange Data Mining. *Media Online)*, *4*(1), 341–349. https://doi.org/10.30865/klik.v4i1.957

Fadlila Nurwanda, Winita Sulandari, Yuliana Susanti, & Zakya Reyhana. (2023). Comparative Analysis Of Performance Levels Of Svm And Naïve Bayes Algorithm For Lifestyle Classification On Twitter Social Media. *INTERNATIONAL CONFERENCE ON DIGITAL ADVANCE TOURISM, MANAGEMENT AND TECHNOLOGY*, *1*(1), 215–230. https://doi.org/10.56910/ictmt.v1i1.65

Fikry, M. (2023). PENDETEKSIAN KALIMAT SINDIRAN MENGGUNAKAN ALGORITMA NAIVE BAYES. *Jurnal Teknologi Terapan & Sains 4*, *0*(3). https://ojs.unimal.ac.id/tts/index

Henderi, Asro, Sulaiman, A., Kurniawan, T. B., Dewi, D. A., & Alqudah, M. (2025). Utilizing Sentiment Analysis for Reflect and Improve Education in Indonesia. *Journal of Applied Data Sciences*, *6*(1), 189–200. https://doi.org/10.47738/jads.v6i1.527

Jannah, N. Z. B., & Kusnawi, K. (2024). Comparison of Naïve Bayes and SVM in Sentiment Analysis of Product Reviews on Marketplaces. *Sinkron*, *8*(2), 727–733. https://doi.org/10.33395/sinkron.v8i2.13559

Kinerja Naïve Bayes, P., Supian, A., Tri Revaldo, B., Marhadi, N., & Efrizoni, L. (n.d.). *Acuan Supian Perbandingan Kinerja Naïve Bayes dan SVM pada Analisis Sentimen Twitter Ibukota Nusantara*. https://github.com/syenirasheila/Sentiment-Analysis-IKN-

Mantik, J., Rasyidin Muqsith Rizqi Prasetyo, I., & Musthafa, A. (2023). Comparison between naive bayes method and support vector machine in sentiment analysis of the relocation of the Indonesian capital. In *Jurnal Mantik* (Vol. 7, Issue 2). Online.

Monica, M., Rukhviyanti, N., Yusron, M., Teknologi Informasi Sekolah Tinggi Manajemen Informatika dan Komputer Indonesia Mandiri, F., & Ekonomi dan Bisnis, F. (2024). *Analisis Literatur Review Perencanaan Strategi Sistem Informasi Menggunakan Metode Pieces Framework Literature Analysis Review of Information System Strategy Planning Using Pieces Framework Method*. https://jurnaldrpm.budiluhur.ac.id/index.php/Kresna/

Putnarubun, D. J., & Palembang, C. F. (n.d.). *PENERAPAN METODE MACHINE LEARNING-NAIVE BAYES PADA ANALISIS SENTIMEN PEMINDAHAN IBU KOTA NEGARA BARU Implementation of Machine Learning-Naïve Bayes Method to Sentiment Analysis of New Country Capitals*.

Rizki Fitrianto, M., & Rizka Sugiarto, A. (2026). Public Perception of IKN as the Capital City of Indonesia After the Budget Blockage Issue through Mention and Sentiment Trend Analysis. *European Public and Social Innovation Review*, *11*. https://doi.org/10.31637/epsir-2026-1903

Robi Padri, A. (2023). HCI DAN MEDIA SOSIAL: STUDI KASUS ANALISIS SENTIMEN PILPRES 2024 DI TWITTER MENGGUNAKAN NAIVE BAYES CLASSIFIER. *Jurnal SIMETRIS*, *14*(2).

Robi Padri, A., Asro, A., & Indra, I. (2023). Classification of Traffic Congestion in Indonesia Using the Naive Bayes Classification Method. *Journal of World Science*, *2*(6), 877–888. https://doi.org/10.58344/jws.v2i6.285

Setiawan, S. B., & Isnain, R. (2024). *JURNAL MEDIA INFORMATIKA BUDIDARMA Sentimen Analisis Masyarakat Terhadap Pembangunan IKN Menggunakan Algoritma Lexicon Based Approach dan Naïve Bayes*. https://doi.org/10.30865/mib.v8i2.7605

Silahudin, D., Henderi, & Holidin, A. (2020). Model expert system for diagnosis of COVID-19 using naïve bayes classifier. *IOP Conference Series: Materials Science and Engineering*, *1007*(1). https://doi.org/10.1088/1757-899X/1007/1/012067

Siregar, A. M. (2023). Analisis Sentimen Pindah Ibu Kota Negara (IKN) Baru pada Twitter Menggunakan Algoritma Naive Bayes dan Support Vector Machine (SVM). *Faktor Exacta*, *16*(3). https://doi.org/10.30998/faktorexacta.v16i3.16703

Sutoyo, E., & Almaarif, A. (2020). Twitter sentiment analysis of the relocation of Indonesia's capital city. *Bulletin of Electrical Engineering and Informatics*, *9*(4), 1620–1630. https://doi.org/10.11591/eei.v9i4.2352

Welly Andi, A., Jumadi, J., Slamet, C., Rialdy Atmadja, A., Sa, D., Maylawati, adillah, & Ali Ramdhani, M. (2025). *Sentiment Analysis of State Capital Relocation of Indonesia using Convolutional Neural Network Methods*. https://orcid.org/0000-0001-7798-8078

Wibawa, A. P., Kurniawan, A. C., Murti, D. M. P., Adiperkasa, R. P., Putra, S. M., Kurniawan, S. A., & Nugraha, Y. R. (2019). Naïve Bayes Classifier for Journal Quartile Classification. *International Journal of Recent Contributions from Engineering, Science & IT (IJES)*, *7*(2), 91. https://doi.org/10.3991/ijes.v7i2.10659

Yanti, I., & Utami, E. (2024). SENTIMENT ANALYSIS OF INDONESIA'S CAPITAL RELOCATION USING WORD2VEC AND LONG SHORT-TERM MEMORY METHOD. *Jurnal Teknik Informatika (Jutif)*, *6*(1), 149–158. https://doi.org/10.52436/1.jutif.2025.6.1.2712

Zamzami, F., Hidayat, R., & Fathonah, R. (2024). PENERAPAN ALGORITMA NAIVE BAYES CLASSIFIER UNTUK ANALISIS SENTIMEN KOMENTAR TWITTER PROYEK PEMBAGUNAN IKN. *Faktor Exacta*, *17*(1). https://doi.org/10.30998/faktorexacta.v17i1.22265