Jurnal Krisnadana (Jurnal Komputer, Sistem Kendali & Jaringan)
Vol 5, Issue 3, May 2026, pp. 614-624
https://doi.org/10.58982/yhck9073

E-ISSN: 2808-3563
P-ISSN: 2808-3520

Sentiment Analysis of the National Capital Relocation to IKN Using TF-IDF and Logistic Regression

Ni Wayan Jeri Kusuma Dewi [1*], Indra Pratistha [2], Tasya Alifah[3], Ni Komang Sri Wahyuni[4]

[1*,2,3,4] Institut Bisnis dan Teknologi Indonesia, Denpasar, Indonesia
[1*]*wayan.kusumadewi@instiki.ac.id*; [2]*indra.pratistha@instiki.ac.id*; [3]*tasyaalifiah97@gmail.com*; [4]*nikomangnopisriwahyuni@gmail.com*

ARTICLE INFO (8 pt)

Article history:
Received 4 June 2026
Revised 16 July 2026
Accepted 17 July 2026

***Keywords**:*
Sentiment Analysis;
TF-IDF;
Logistic Regression;
Emolex;

ABSTRACT

The relocation of the national capital from Jakarta to East Kalimantan has sparked various reactions from the public, many of which have been expressed through social media platforms such as YouTube. This study aims to analyze public responses from the perspective of sentiment and emotion toward the policy. The methodology employed involves keyword weighting using TF-IDF and Logistic Regression techniques. The data was then processed through preprocessing stages and labeled using the EmoLex Dictionary. Out of a total of 5,835 comments analyzed, the results showed that 2,180 comments (37.4%) were neutral in sentiment, followed by 2,077 comments (35.6%) with positive sentiment and 1,578 comments (27.0%) containing negative sentiment. In emotion classification, anger was the most dominant emotion, with 2,525 comments (43.3%), followed by trust with 897 comments (15.4%), anticipation with 750 comments (12.9%), and other emotions, such as disgust, neutral, sadness, fear, joy, and surprise, with smaller proportions. The results of the model performance evaluation for sentiment classification showed an accuracy of 93,9%, a precision of 94,6%, a recall of 93,0%, and an F1-score of 93,5%.

1. Introduction

The relocation of the capital city from Jakarta to Ibu Kota Nusantara (IKN) in East Kalimantan has been a major decision that sparked diverse public reactions. While viewed as an effort to alleviate congestion in Jakarta, the move has also raised concerns regarding social, economic, and environmental impacts [1]. The public has voiced opinions ranging from support to criticism via social media. Given the proliferation of discussions and comments on digital platforms, sentiment analysis is essential to objectively identify patterns in public opinion [2]. Such analysis can provide insights into how the public perceives the policy and the potential social impacts on their daily lives [3],[4].

YouTube was selected for this study due to its large user base and its frequent use by the public as a source of information [5],[6]. According to a report by We Are Social (2024), YouTube is the social media platform with the highest number of users in Indonesia, reaching over 139 million people. YouTube comments allow users to express their opinions in greater detail compared to other social media platforms. Furthermore, YouTube serves as a forum for discussion among diverse groups including academics, journalists, and the general public who wish to share information and perspectives [7]. Consequently, this study utilizes comments from videos discussing the relocation of the capital city as the primary data for sentiment analysis.

The Term Frequency-Inverse Document Frequency (TF-IDF) method and the Logistic Regression algorithm are employed in this study due to their strengths in text processing and sentiment classification. TF-IDF transforms text into a numerical representation by accounting for a word's frequency within a document and its significance across the entire dataset [8],[9]. This method is highly effective at identifying words that carry substantial informational weight for sentiment analysis. Meanwhile, Logistic Regression is a widely used probability-based classification algorithm that demonstrates strong performance across various text analysis tasks [10],[11],[12].

Various previous studies have demonstrated that Logistic Regression often yields higher accuracy in text-based sentiment analysis. For instance, research by Ramadhani and Suryono (2024) titled "Comparison of Naive Bayes and Logistic Regression Algorithms for Metaverse Sentiment Analysis" showed that the Logistic Regression method achieved 95% accuracy, surpassing the Naive Bayes method, which attained 91%. Furthermore, a study by Budianto et al. (2024) revealed that Logistic Regression outperformed other methods in identifying positive sentiment, achieving an F1-score of 0.89 higher than that of SVM (0.78) along with fewer misclassifications (110) and higher True Positive and True Negative values. Consequently, combining the TF-IDF method with the Logistic Regression algorithm represents an appropriate choice for analyzing sentiment regarding the capital city relocation based on YouTube comments.

The relocation of the capital city affects various segments of society in Jakarta, East Kalimantan, and other regions. Residents of Jakarta may experience changes in the economic and administrative sectors as the seat of government shifts [13]. Meanwhile, residents of East Kalimantan will face social and environmental transformations resulting from massive development projects in the area. Furthermore, communities outside these two regions may also experience indirect impacts, such as shifts in labor migration patterns and investment flows [14].

Therefore, the author conducted a study titled “Sentiment Analysis of the National Capital Relocation to IKN on YouTube Using TF-IDF and Logistic Regression,” which aims to identify and understand how people from various regions respond to the capital city relocation and to assess the extent to which this policy impacts their social lives both positively and negatively.

2. Method

2.1 Crawling Data

The sentiment analysis in this study began with the process of crawling user reviews from YouTube regarding the relocation of the National Capital (IKN) using Python. The data collection, covering the period from January 2024 to December 2024, yielded 5,851 data entries. Data retrieval was conducted using Google Colab to execute Python code, utilizing an API Key (YouTube Data API v3) from the Google Cloud Console.

2.2 Preprocessing Data

Data preprocessing is the process of preparing and transforming raw data before it is used for analysis [15],[16]. This stage is crucial because data obtained through crawling is often not yet ready to yield accurate sentiment analysis. Several steps are carried out during this phase, including cleansing, case folding, tokenizing, and stopword removal, as illustrated in Figure 1 below:

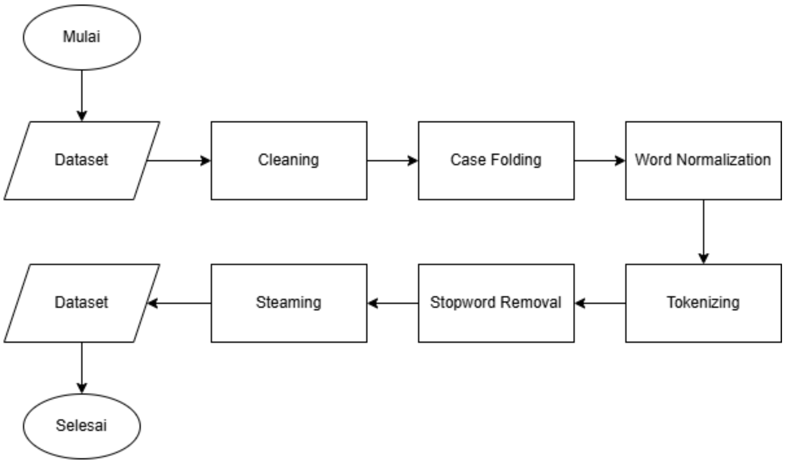

Figure 1. Preprocessing
[Source: Darmawan et al., 2023]

2.3 TF-IDF Weighting

Following preprocessing, a value or weight is assigned to each word using TF-IDF. Word weighting is the process of assigning a value or weight to every word found within a document. In the context of sentiment analysis, each word is assigned a specific value; the higher the value, the more frequently the word appears in the sentence or document. The TF-IDF calculation is as described in the following equation:

1. The Term Frequency (TF) value is obtained from the frequency of occurrence of feature t in document d.
   $$TF_t = (t,d) \tag{1}$$
2. The Inverse Document Frequency value is obtained from the logarithm of the total number of documents (n) divided by the number of documents (df) containing feature t.
   $$IDF_t = log\frac{n}{df(t)} \tag{2}$$
3. The Term Frequency-Inverse Document Frequency (Wt) value is obtained by multiplying the TF value by the IDF value.
   $$W_t = TF_t\ x\ IDF_t \tag{3}$$

2.4 Data Partitioning

The labeled data is divided into two parts: training data (80%) and test data (20%). This split aims to train the model and evaluate its accuracy. By using 80% of the data for training, the model has a larger number of samples from which to learn existing patterns and characteristics. This potentially enhances accuracy when the model encounters previously unseen test data. Following training, the labeled test data is used to evaluate the model's performance using metrics such as precision, recall, accuracy, and F1-score. Of the 5,835 data points, 4,668 were allocated for training and 1,167 for testing.

2.5 Logistic Regression Model Training

During the training phase, Logistic Regression learns the relationship between data features and the target class by constructing a mathematical model based on the logistic (sigmoid) function. The model aims to calculate the probability of a data point belonging to a specific class by determining the optimal coefficients (weights) that minimize prediction error [17],[18]. Once the model is established, a testing phase is conducted to evaluate its accuracy using previously unseen data. The results of this testing provide evaluation metrics such as True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) that help assess the model's performance. The basic formula for the sigmoid function in logistic regression is as follows:

$$P(y = 1|X) = 1 / 1 + e^{-(\beta 0 + \beta 1 X1 + \beta 2 X2 + \cdots + \beta n Xn)} \quad (4)$$

Description of Equations:

X	: Input feature vector
y	: Target label or class (e.g., 0 or 1)
β 0	: Intercept (bias)
β 1, …, β n	: Regression coefficients for each feature
e	: Euler's number (approximately 2.718)

2.6 Dataset Label Changes

Test data labels, initially designated as positive and negative, were converted into binary values specifically zero (0) and one (1) to facilitate model evaluation using a confusion matrix. This binary representation simplifies the identification of the confusion matrix's key components and clarifies the model's classification evaluation process. Furthermore, this conversion streamlines the calculation of evaluation metrics such as precision, recall, accuracy, and F1-Score. The formulas for precision, accuracy, and recall are as follows:

1. $$Precision = \frac{TP}{TP + FP} \times 100\% \quad (5)$$
2. $$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad (6)$$
3. $$Recall = \frac{TP}{TP + FN} \times 100\% \quad (7)$$
4. $$F1 - Score = 2x\ \frac{Precision\ x\ Recall}{Precision + Recall} \times 100\% \quad (8)$$

Description of Equations:
True Positive (TP), Predicting a positive value, and the prediction is correct.
False Positive (FP), Predicting a positive value, and the prediction is incorrect.
True Negative (TN), Predicting a negative value, and the prediction is correct.

False Negative (FN), Predicting a negative value, and the prediction is incorrect.

Table 1. Confusion Matrix
[Source: Supono & Suprayogi, 2021].

Confusion Matrix		True Value	
		Classified as Positive	Classified as Negative
Predicted Value	Positive	TP (True Positif) Correct Result	FP (False Positive) Uneexpected Result
	Negative	FN (False Negative) Missing Result	TN (True Negative) Correct Absence of Result

3. Result and Discussions

3.1 Data Decription

The data used in this study consists of YouTube user comments regarding the relocation of the National Capital (IKN). Data was collected using the YouTube Data API v3 through a crawling process with the keyword "IKN" covering the period from January 2024 to December 2024. The data retrieval process yielded 5,851 comments; after removing duplicates, 5,835 comments remained and were used as the research dataset. Prior to the classification process, all data undergoes text preprocessing stages, which include: Cleaning, Case Folding, Word Normalization, Tokenizing, Stopword Removal, Stemming. This stage aims to clean the text of various forms of noise, standardize word forms, and produce a dataset ready for TF-IDF weighting and classification using the Logistic Regression algorithm. Subsequently, a labeling process is carried out using the NRC Emotion Lexicon (EmoLex), assigning each comment a sentiment label (positive, negative, or neutral) and a dominant emotion label (anger, anticipation, disgust, fear, joy, sadness, surprise, trust, or neutral).

3.2 Model Implementation

The research was implemented using the Python programming language on the Google Colaboratory platform, utilizing libraries such as Pandas, NumPy, Scikit-learn, NLTK, Sastrawi, Matplotlib, and WordCloud.

1. Explonatory Data Analysis

   At this stage, data that has undergone text preprocessing or text cleaning is explored. The goal of Exploratory Data Analysis (EDA) is to provide an overview of the data; this is achieved through methods such as generating word clouds or identifying the most frequently occurring words within the cleaned collection of YouTube comments regarding the IKN. The first step is to import the necessary libraries, such as pandas, wordcloud, and matplotlib [19]. Missing values in the comments column are filled with empty strings to prevent errors. All comments are then combined into a single block of text. The code also defines "stopwords" common, irrelevant words to ensure they are excluded from the visualization. A WordCloud object is created with a black background, a maximum of 700 words, and dimensions of 1600x800 pixels. Once the WordCloud is generated from the text, the result is displayed using matplotlib. The resulting WordCloud appears as follows:

Figure 2. Word Cloud Results
[Source: the author themselves]

Based on the Word Cloud visualization shown above, several words appear with high frequency and are displayed prominently. Words such as "IKN," "Jokowi," "bangun" (build/construction), "rakyat" (people/citizens), and "Indonesia" stand out, indicating that the main topics in the comments are closely linked to the Ibu Kota Nusantara (IKN) development project and the role of President Jokowi. The words "bangun" and "kota" (city) reflect the public's focus on physical development, while "rakyat," "bangga" (proud), and "hidup" (life) suggest emotional engagement or public expectations regarding the changes brought about by the project. Furthermore, the words "Kalimantan," "Jakarta," and "presiden" clarify the geographical and political context of the discussion. The visual dominance of these words demonstrates that the IKN project is a focal point of public attention and highlights the interplay between the government, the public, and the vision for national development. This Word Cloud provides a quick overview of the dominant themes emerging from the analyzed collection of comments.

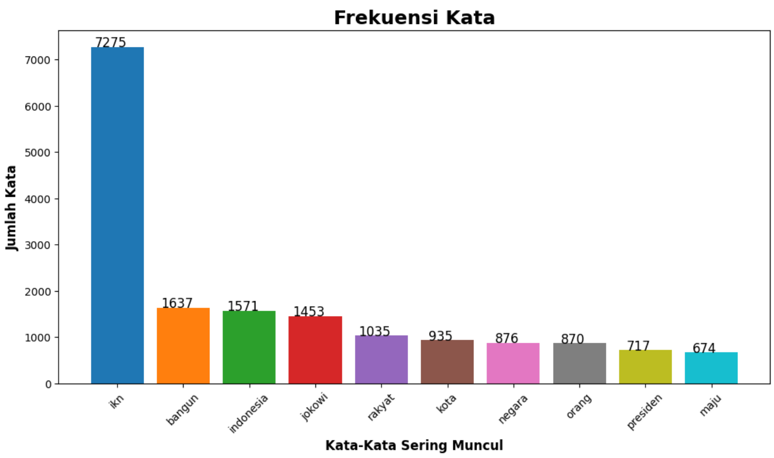

Figure 3. Word Frequency
[Source: the author themselves]

The implementation stages included data crawling, text preprocessing, TF-IDF weighting, splitting the data into training and testing sets, training a Logistic Regression model, evaluation using a confusion matrix, and visualization of the analysis results. During the weighting stage, the TF-IDF (Term Frequency–Inverse Document Frequency) method was employed to convert text data into a numerical representation suitable for processing by the Logistic Regression algorithm. Subsequently, the dataset was divided into 80% training data (4,668 samples) and 20% testing data (1,167 samples) prior to model training. The Logistic Regression model was then trained on the training data to learn the patterns of association between TF-IDF weights and the sentiment labels obtained through an automated labeling process. The trained model was subsequently used to make predictions on the testing data and was evaluated using a confusion matrix, along with metrics for Accuracy, Precision, Recall, and F1-Score.

3.3 Discussion

The research results demonstrate that combining TF-IDF with Logistic Regression yields a high-accuracy sentiment classification model following parameter tuning. TF-IDF weighting successfully identifies key terms within YouTube comments, ensuring the information fed into the model is more representative than that derived from simple word frequency alone. The sentiment distribution reveals that the majority of public comments regarding the National Capital City (IKN) relocation are neutral (37.4%), followed by positive sentiment (35.6%) and negative sentiment (27.0%). This indicates that most of the public is adopting a "wait-and-see" approach regarding the IKN project's progress, while others express either support for or opposition to the policy. Overall, the study confirms that the combination of TF-IDF and Logistic Regression achieves excellent classification performance after tuning. The resulting model can serve as a tool to gauge public opinion on the capital city relocation policy based on comments posted on YouTube.

1. Result Comparasion

   Model evaluation was conducted before and after the hyperparameter tuning process to determine the impact of optimization on the Logistic Regression model's performance. The sentiment classification evaluation results show that prior to tuning, the model achieved an Accuracy of 63.4%, Precision of 66.8%, Recall of 62.7%, and F1-Score of 62.6%. Following tuning, all metrics improved significantly, reaching an Accuracy of 93.3%, Precision of 94.6%, Recall of 93.0%, and F1-Score of 93.5%. The following table shows a comparison of model performance before and after tuning.

Table 2. Comparison of model performance before and after tuning.
[Source: the author themselves].

Metric	Before Tuning	After Tuning
Accuracy Sentimen	63,40%	93,30%
Precision Sentimen	66,80%	94,60%
Recall Sentimen	62,70%	93,00%

F1-Score Sentimen	62,60%	93,50%

2. Sentiment Classification
   Categorizing sentiment data into positive, negative, and neutral groups is a crucial step in sentiment analysis for gaining a more comprehensive understanding of public opinion [20]. By classifying data into these three sentiment types, analysts can identify the general trend regarding a topic whether the prevailing stance is supportive, opposed, or neutral. This combination of sentiments greatly facilitates a richer and more meaningful interpretation of the data. The sentiment pie chart visualization makes it easy to understand the distribution of data based on sentiment within the dataset.

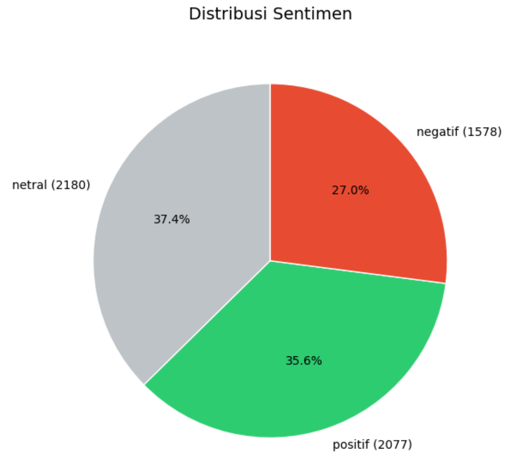

Figure 4. Sentiment Pie Chart
[Source: the author themselves]

Based on the pie chart showing the sentiment distribution regarding the relocation of the National Capital (IKN), the majority of public sentiment is neutral, accounting for 2,180 data points (37.4%). This indicates that most of the public has not yet adopted a clear stance on the relocation issue; people tend to be cautious, await further developments, or feel they lack sufficient information to form a definitive opinion. Meanwhile, positive sentiment accounts for 2,077 data points (35.6%), reflecting views that support and are optimistic about the capital relocation policy. This support likely stems from hopes for more equitable development, the alleviation of the burden on Jakarta, and the potential for the growth of new economic hubs outside Java. Meanwhile, negative sentiment accounted for 1,578 data points (27.0%), indicating public concern or disapproval regarding this policy. These concerns may relate to aspects such as the state budget, infrastructure readiness, environmental impact, or potential social inequality in the new region. This distribution reveals that while many members of the public support or hold a positive view of the relocation of the National Capital (IKN), a significant portion remains neutral, and the remainder express objections. Understanding this distribution is crucial for the government and policymakers to design more inclusive and informative public communication strategies and to openly address emerging concerns. Effective efforts to engage the neutral group and respond to negative sentiment with constructive solutions could be key to building legitimacy and public support for the strategic agenda of relocating the national capital.

3. Neutral Sentiment

   Neutral sentiment is dominated by words such as "IKN," "Jokowi," "build," "Indonesia," "state," "city," "people," "project," and "work," indicating that the majority of the public discusses the issue of the National Capital relocation factually, without strong emotional bias. The appearance of words like "hope," "see," and "think" reflects an observational stance, while terms such as "thank you," "cool," and "stalled" reveal a diversity of opinions. This suggests that a portion of the public is still awaiting further developments regarding the IKN project before forming a definitive stance.

Figure 5. Neutral Sentiment Word Cloud

[Source: the author themselves]

4. Positive Sentiment

   Positive sentiment is dominated by words such as "IKN," "build," "Jokowi," "Indonesia," "nation," "advanced," "work," "president," and "people," reflecting public optimism regarding the development of the IKN and Indonesia's future. Additionally, words like "support," "thank you," "enthusiasm," "cool," and "good" reflect appreciation, support, and positive expectations concerning the capital city relocation project.

Figure 6. Positive Sentiment Word Cloud

[Source: the author themselves]

5. Negative Sentiment

   Negative sentiment is dominated by words such as "IKN," "Jokowi," "build," "state," "people," "project," "government," and "Indonesia," reflecting public dissatisfaction with the policy to relocate the national capital and the government's performance. Furthermore, words like "corruption," "failed," "rotten," "wrong," and "step down" indicate criticism regarding development transparency, budget utilization, and the social impact of the IKN project.

Figure 7. Negative Sentiment Word Cloud
[Source: the author themselves]

4. Conclusion

Based on the analysis of sentiment distribution regarding the relocation of the National Capital (IKN) and the performance evaluation of the proposed classification model, it can be concluded that the sentiment analysis framework, which integrates data crawling, text preprocessing, EmoLex-based sentiment labeling, TF-IDF feature extraction, and Logistic Regression classification, demonstrates strong effectiveness in analyzing public opinion expressed in YouTube comments. The classification model achieved an accuracy of 93.9%, with a precision of 94.6%, recall of 93.0%, and an F1-score of 93.5%, indicating that the proposed approach is highly reliable for classifying public sentiment toward the IKN issue. Furthermore, the sentiment distribution analysis shows that neutral sentiment dominates the discussion, accounting for 2,180 comments (37.4%), followed by positive sentiment with 2,077 comments (35.6%) and negative sentiment with 1,578 comments (27.0%). These findings suggest that while public opinion includes both support and criticism regarding the relocation of the national capital, a considerable proportion of users remain neutral, reflecting a tendency to await further developments or maintain a cautious perspective on the issue.

For future research, it is recommended to explore more advanced deep learning approaches, such as Bidirectional Encoder Representations from Transformers (BERT) or other transformer-based models, to improve sentiment classification performance and capture contextual information more effectively. Future studies may also compare multiple sentiment lexicons or labeling strategies, incorporate comments from various social media platforms to improve data diversity and generalizability, and conduct temporal sentiment analysis to examine how public opinion evolves over time. In addition, rather than relying primarily on word cloud visualization, future research should employ more informative visualization techniques, such as topic modeling, keyword co-occurrence networks, semantic relationship graphs, or interactive dashboards, to provide deeper insights into public discourse and the factors influencing sentiment toward the IKN policy.

Acknowledgment

The author expresses gratitude to God Almighty for His grace and blessings, which enabled the successful completion of this research. The author also extends thanks to the academic supervisor for their guidance, direction, and invaluable input throughout the research process. Appreciation is also expressed to the Institut Bisnis dan Teknologi Indonesia for the academic support and facilities provided. Furthermore, the author thanks all those who assisted whether directly or indirectly in the successful completion of this research.

References

[1] S. J. Purnama and Chotib, "Analisis Kelayakan Pendirian Usaha Pengolahan Limbah Medis untuk Meningkatkan Pendapatan Asli Daerah," *J. Ekon. dan Kebijak. Publik*, vol. 13, no. 1, pp. 57–70, 2022, doi: 10.22212/jekp.v13i1.2155.

[2] D. Duei Putri, G. F. Nama, and W. E. Sulistiono, "Analisis Sentimen Kinerja Dewan Perwakilan Rakyat (DPR) Pada Twitter Menggunakan Metode Naive Bayes Classifier," *J. Inform. Dan Tek. Elektro Terap.*, vol. 10, no. 1, pp. 34–40, 2022, doi: 10.23960/jitet.v10i1.2262.

[3] M. Sabiq, S. To Anwar, S. Muhammad, Hasbi, and Arisnawawi, "Perubahan Sosial Masyarakat Pedalaman (Studi Masyarakat Adat Kalimantan Timur Pada Proses Pemindahan Ibu Kota Negara)," in *Konferensi Nasional Sosiologi IX APSSI 2022 Balikpapan*, 2022, pp. 1–3. [Online]. Available: https://pkns.portalapssi.id/index.php/pkns/article/view/17

[4] G. Darmawan, S. Alam, and M. I. Sulistyo, "Analisis Sentimen Berdasarkan Ulasan Pengguna Aplikasi Mypertamina Pada Google Playstore Menggunakan Metode Naïve Bayes," *STORAGE – J. Ilm. Tek. Dan Ilmu Komput.*, vol. 2, no. 3, pp. 100–108, 2023.

[5] Mahpudoh and D. Romdhoningsih, "Pemanfaatan Sosial Media YouTube Sebagai Media Pembelajaran Bahasa Indonesia di Perguruan Tinggi," *J. Kiprah Pendidik.*, vol. 9, no. 2, pp. 109–114, 2024.

[6] A. P. Giovani, A. Ardiansyah, T. Haryanti, L. Kurniawati, and W. Gata, "Analisis Sentimen Aplikasi Ruang Guru Di Twitter Menggunakan Algoritma Klasifikasi," *J. Teknoinfo*, vol. 14, no. 2, p. 115, 2020, doi: 10.33365/jti.v14i2.679.

[7] J. P. Suharsono and D. Nurahman, "Pemanfaatan Youtube Sebagai Media Peningkatan Pelayanan Dan Informasi," *Ganaya J. Ilmu Sos. Dan Hum.*, vol. 7, no. 1, pp. 298–304, 2024, doi: 10.37329/ganaya.v7i1.3157.

[8] G. H. A. H. Noer, "Implementasi Algoritma Naive Bayes dan TF-IDF dalam Analisis Sentimen Data Ulasan (Studi Kasus: Ulasan Review Aplikasi E-Commerce Shopee Di Situs Google Playstore)," 2024.

[9] R. A. Supono and M. A. Suprayogi, "Perbandingan Metode TF-ABS dan TF-IDF Pada Klasifikasi Teks Helpdesk Menggunakan K-Nearest Neighbor," *J. RESTI (Rekayasa Sist. Dan Teknol. Informasi)*, vol. 5, no. 5, pp. 911–918, 2021, doi: 10.29207/resti.v5i5.3403.

[10] S. A. Assaidi and F. Amin, "Analisis Sentimen Evaluasi Pembelajaran Tatap Muka 100 Persen pada Pengguna Twitter menggunakan Metode Logistic Regression," *J. Pendidik. Tambusai*, vol. 6, no. 2, pp. 13217–13227, 2022.

[11] B. Ramadhani and R. R. Suryono, "Komparasi Algoritma Naïve Bayes dan Logistic Regression Untuk Analisis Sentimen Metaverse," *J. Media Inform. Budidarma*, vol. 8, no. 2, p. 714, 2024, doi: 10.30865/mib.v8i2.7458.

[12] A. G. Budianto, R. Rusilawati, A. T. E. Suryo, G. R. Cahyono, A. F. Zulkarnain, and M. Martunus, "Perbandingan Performa Algoritma Support Vector Machine (SVM) dan Logistic Regression untuk Analisis Sentimen Pengguna Aplikasi Retail di Android," *J. Sains Dan Inform.*, vol. 10, no. 2, pp. 1–10, 2024, doi: 10.34128/jsi.v10i2.911.

[13] A. Siti, A. N. Aditya, S. Ahmad, and R. Farida, "Analisis Dampak Dan Resira Terhadap Ekonomi Di Indonesia," *J. Ekon.*, vol. 2, no. 2, pp. 10–18, 2023.

[14] I. Irwansyah, D. Daryono, and I. J. Mangele, "Dampak Pemindahan Ibukota Ke Kalimantan Timur : Perspektif Democracy Civil Society," *Innov. J. Soc. Sci. Res.*, vol. 4, no. 3, pp. 6681–6691, 2024.

[15] M. R. S. Alfarizi, M. Al-farish, M. Z. Taufiqurrahman, G. Ardiansah, and M. Elgar, "Penggunaan Python Sebagai Bahasa Pemrograman untuk Machine Learning dan Deep Learning," *Karya Ilm. Mhs. Bertauhid (KARIMAH TAUHID)*, vol. 2, no. 1, pp. 1–6, 2023.

[16] J. E. Widyaya and S. Budi, "The Effect of Preprocessing on the Classification of Diabetic Retinopathy with the Transfer Learning Convolutional Neural Network Approach," *J. Tek. Inform. Dan Sist. Inf.*, vol. 7, no. 1, pp. 110–124, 2021.

[17] M. Cindo and D. P. Rini, "Literatur Review: Metode Klasifikasi Pada Sentimen Analisis," in *Seminar Nasional Teknologi Komputer & Sains (SAINTEKS)*, 2019, pp. 66–70. [Online]. Available: https://seminar-id.com/semnas-sainteks2019.html

[18] A. K. N. Bany, "Analisis Sentimen dan Deteksi Emosi dengan Pendekatan Lexicon pada Judul Berita Media Online Mengenai COVID-19 di Indonesia," 2022.

[19] K. Septiani, "Perbandingan Analisis Sentimen Terhadap Pembayaran Digital ' Go - Pay ' Dan ' Ovo ' Di Media Sosial Twitter Menggunakan Metode Naive Bayes Dan Word Cloud," *J. Sos. dan Teknol.*, vol. 2, no. 10, pp. 1–10, 2022.

[20] P. S. Hasugian and J. R. Sagala, "Penerapan Data Mining Untuk Pengelompokan Siswa Berdasarkan Nilai Akademik dengan Algoritma K-Means," *KLIK Kaji. Ilm. Inform. Dan Komput.*, vol. 3, no. 3, pp. 262–268, 2022, [Online]. Available: https://djournals.com/klik