XGBOOST-BASED SENTIMENT ANALYSIS OF MACROECONOMIC POLICY YOUTUBE COMMENTS

Aprilisa Arum Sari[1*], Cindy Nasywa Aurora[2], Jashila Nair Mogan[3]
Universitas Duta Bangsa Surakarta[1,2], Institution[2], Institution[3]
*Correspondence Email : aprilisa_arumsari@udb.ac.id

ABSTRACT

Public reaction to macroeconomic policy is increasingly expressed through YouTube comment sections, yet this unstructured text is rarely used to gauge sentiment toward government decisions such as exchange-rate management. This study builds and evaluates an XGBoost-based classifier to automatically label the sentiment of YouTube comments discussing Indonesia's macroeconomic policy, particularly the depreciation of the Rupiah. A dataset of 17,838 manually labeled Indonesian-language comments (positive, negative, neutral) was collected and preprocessed through case folding, text cleaning, Indonesian stopword removal, and stemming using the Sastrawi library, then represented as 5,000-dimensional TF-IDF feature vectors. An XGBoost multi-class classifier was trained on an 80:20 stratified train-test split. The model achieved an overall accuracy of 85.68%, with the highest F1-score obtained for the neutral class (0.90) and a comparatively lower recall for the positive class (0.70) due to class imbalance. Word cloud visualization revealed that terms such as "pemerintah," "rupiah," and "ekonomi" dominated across all sentiment classes, while "melemah" and "krisis" distinctly marked negative sentiment. These findings demonstrate that XGBoost combined with TF-IDF is an effective and computationally efficient approach for monitoring public sentiment toward macroeconomic policy on social media.

KEYWORDS sentiment analysis, XGBoost, TF-IDF, YouTube comments, macroeconomic policy

This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International

INTRODUCTION

The rapid growth of digital technology has transformed social media into a major platform for public discourse on socioeconomic issues, including macroeconomic policies

implemented by governments(Mohapatra et al., 2022; Tangkudung et al., 2024). Among various social media platforms, YouTube has become one of the largest sources of user-generated content in Indonesia, functioning not only as an entertainment platform but also as a space where users actively express opinions through comment sections(Kemp, 2025; Ma'aly et al., 2024; Mandhasiya et al., 2024). Videos discussing economic issues, such as the depreciation of the Indonesian Rupiah against the US Dollar, often attract thousands of comments reflecting public concerns, criticism, and support toward government policies(Guo et al., 2024; Ma'aly et al., 2024). This large volume of textual data provides valuable information for understanding public perception but is difficult to analyze manually due to its scale and unstructured nature(sabbah Naiarni panjaitana, Azzikrarairyahannum, Syahilatuzzahira, 2024; Setyawan et al., 2025).

Sentiment analysis, a branch of text mining, has been widely adopted to automatically identify the polarity of opinions expressed in textual data by categorizing them into positive, neutral, or negative sentiments (Adilla et al., 2024; Angdresey et al., 2025). It has been successfully applied in product review analysis, political communication, and public policy evaluation because it enables large collections of unstructured text to be transformed into meaningful information for decision-makers(Alnofeli et al., 2026). In the context of macroeconomic policy, sentiment analysis is particularly important because public opinion may influence institutional trust, economic expectations, and market stability(Le Dong Xuan et al., 2026; Patil et al., 2023). Consequently, automated sentiment classification offers a practical alternative to conventional survey-based approaches for monitoring public responses to economic policy.

Among conventional machine learning algorithms, Extreme Gradient Boosting (XGBoost) has demonstrated competitive performance in sentiment classification due to its ability to handle high-dimensional sparse feature representations generated by Term Frequency–Inverse Document Frequency (TF-IDF)(Moghadamnejad et al., 2026; Pardede et al., 2026; Pratiwi et al., 2025). Based on the gradient boosting framework, XGBoost constructs an ensemble of decision trees sequentially, where each tree minimizes the prediction errors of its predecessors, resulting in high predictive performance while reducing the risk of overfitting. These characteristics make XGBoost particularly suitable for sentiment classification on large-scale social media datasets.(Samih et al., 2023; Tsiourlini et al., 2024)

Previous studies have demonstrated the effectiveness of XGBoost for sentiment classification across various domains. Hama Aziz & Dimililer (2021) reported that an XGBoost-based ensemble combining bag-of-words, TF-IDF, and lexicon-based features outperformed individual classifiers in social media sentiment analysis. Similarly, Samih et al., (2023) showed that XGBoost combined with word embeddings achieved superior performance compared with conventional TF-IDF and Doc2Vec models. Studies focusing on YouTube comments by Shivsharan et al., (2025) and Le et al., (2024) further confirmed that XGBoost performs effectively on noisy and unstructured user-generated text. In Indonesia, however, sentiment analysis of macroeconomic discussions remains relatively limited. Existing studies have primarily focused on X (formerly Twitter) and employed conventional classifiers such as Support Vector Machine (SVM). For example, Hisyam achieved 78% classification accuracy using LinearSVC with SMOTE on Indonesian tweets discussing Rupiah depreciation, while Hadi Arfian et al., (2025) and Mustofa & Novita, (2022) also utilized SVM-based approaches for public policy analysis.

Despite these advances, three important research gaps remain. First, previous Indonesian-language studies have largely concentrated on Twitter, whereas large-scale YouTube comments discussing macroeconomic policy have received limited attention. Second, the effectiveness of XGBoost for Indonesian-language YouTube sentiment

classification on macroeconomic discussions has not been systematically investigated. Third, few studies have examined public sentiment specifically related to Indonesia's macroeconomic policy and Rupiah exchange rate depreciation, despite their importance for economic policymaking.

To address these gaps, this study develops and evaluates an XGBoost-based sentiment classification model using 17,838 Indonesian-language YouTube comments discussing Indonesia's macroeconomic policy, with particular emphasis on Rupiah exchange rate depreciation. The proposed approach applies Indonesian-language text preprocessing using Sastrawi stemming and stopword removal, followed by TF-IDF feature extraction and multiclass sentiment classification. In addition to evaluating the model quantitatively, this study employs word cloud visualization to provide qualitative insights into the dominant topics and linguistic characteristics associated with each sentiment category.

The main contributions of this study are threefold. First, it provides one of the first large-scale investigations of Indonesian-language YouTube comments related to macroeconomic policy. Second, it empirically evaluates the effectiveness of XGBoost combined with TF-IDF for multiclass sentiment classification in this domain. Third, it demonstrates how automated sentiment analysis can support timely, data-driven monitoring of public opinion toward macroeconomic policies, providing useful insights for policymakers, researchers, and government institutions.

RESEARCH METHOD

This study adopts the Cross-Industry Standard Process for Data Mining (CRISP-DM) as the research methodology because it provides a systematic and structured framework for developing machine learning models(A. A. Sari & Sari, 2026; O. K. Sari & Cherid, 2023). The methodology consists of six phases: Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation, and Deployment. Figure 1 illustrates the overall research framework.

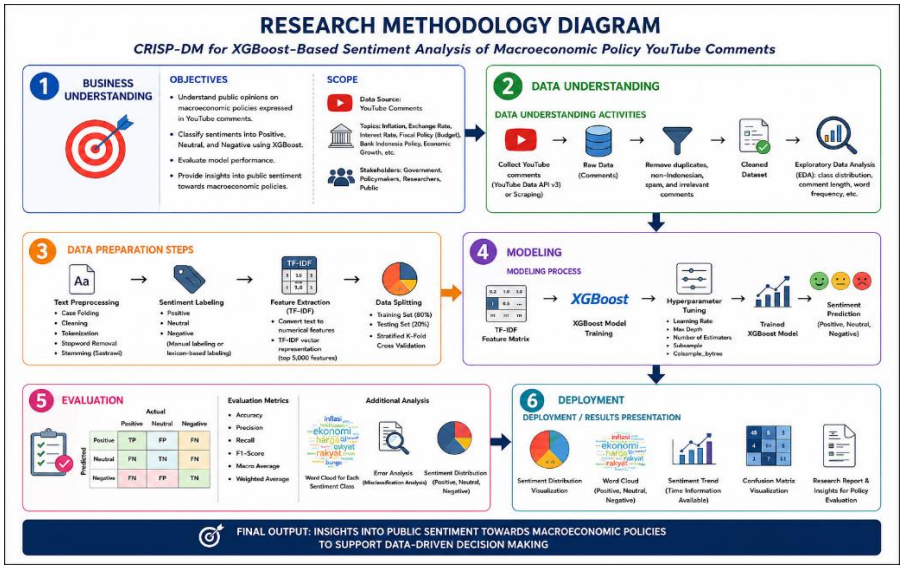


Figure 1. Research Methodology Diagram

In the Business Understanding phase, this study aims to analyze public sentiment toward Indonesia's macroeconomic policies, including inflation, exchange rate

fluctuations, interest rate decisions, fiscal policy, and Bank Indonesia's monetary policy, as reflected in YouTube comments. The primary objective is to develop an XGBoost-based sentiment classification model capable of categorizing comments into three sentiment classes: positive, neutral, and negative. Furthermore, the study evaluates the classification performance of the proposed model using standard multiclass evaluation metrics and derives insights into public perceptions of macroeconomic policies. These insights are expected to support evidence-based policy evaluation and assist policymakers, government institutions, researchers, and other stakeholders in understanding public responses to macroeconomic issues.

In the Data Understanding phase, the pre-labeled dataset (labeled_comments.csv) was imported into a Pandas DataFrame for initial exploration. The dataset consists of raw and preprocessed Indonesian-language YouTube comments, each annotated with one of three sentiment labels: positive, neutral, or negative. An initial data exploration was conducted to examine the dataset structure, attribute types, and data completeness using descriptive statistical functions. This process also included verifying the consistency of sentiment labels and identifying potential data quality issues, such as missing or duplicated records. These exploratory analyses provided a comprehensive understanding of the dataset and ensured its suitability for subsequent preprocessing and model development stages.

In the Data Preparation phase, the textual data were transformed into a machine-learning-ready format through a sequence of preprocessing steps. Initially, all comments were converted to lowercase (case folding) and cleaned using regular expressions to remove URLs, punctuation, numbers, special characters, and other non-alphabetic symbols while preserving whitespace. The cleaned text was subsequently tokenized into individual words. To reduce noise and eliminate terms with limited semantic contribution, Indonesian stopwords obtained from the Natural Language Toolkit (NLTK) corpus were removed. The remaining tokens were then normalized through stemming using the Sastrawi library to convert inflected words into their root forms. Following text preprocessing, the documents were represented numerically using the Term Frequency–Inverse Document Frequency (TF-IDF) technique, with the vocabulary limited to the 5,000 most informative features to reduce feature dimensionality while retaining discriminative terms. The sentiment labels were encoded into numerical values to facilitate supervised learning. Finally, the processed dataset was partitioned into training and testing subsets using an 80:20 stratified split with random_state = 42, ensuring that the original sentiment class distribution was preserved in both subsets and enabling a reliable evaluation of the classification model.

In the Modeling phase, the Extreme Gradient Boosting (XGBoost) algorithm was employed to develop the sentiment classification model. XGBoost was selected because of its computational efficiency, scalability, and ability to handle high-dimensional sparse feature representations generated by TF-IDF. Since the task involves classifying comments into three sentiment categories (positive, neutral, and negative), the model was configured as a multiclass classification problem using the multi:softmax objective function, while multiclass logarithmic loss (mlogloss) was employed as the evaluation metric during training. The TF-IDF feature matrix obtained from the training dataset served as the model input, enabling XGBoost to iteratively construct an ensemble of decision trees by minimizing the classification error through gradient boosting. After the training process was completed, the learned model was used to predict the sentiment labels of the unseen testing dataset for subsequent performance evaluation.

In the Evaluation phase, the trained XGBoost model was evaluated using the unseen testing dataset to assess its ability to classify YouTube comments into positive, neutral, and negative sentiment categories. Model performance was measured using standard multiclass classification metrics, including Accuracy, Precision, Recall, and F1-score, which provide

a comprehensive assessment of overall and class-specific performance. In addition, a confusion matrix was generated to analyze the distribution of correctly classified instances and misclassifications across sentiment classes, allowing a more detailed examination of the model's classification behavior. To complement the quantitative evaluation, word cloud visualizations were created for each sentiment category to identify the most frequently occurring terms and facilitate qualitative interpretation of linguistic characteristics associated with different public sentiments. The quantitative and qualitative evaluation results are presented and discussed in the Results and Discussion section.

In the Deployment phase, the developed XGBoost sentiment classification model is intended to support automated monitoring of public opinion toward Indonesia's macroeconomic policies using YouTube comments. By classifying comments into positive, neutral, and negative sentiments, the model can assist policymakers, government institutions, and researchers in identifying public responses to monetary and fiscal policy announcements more efficiently than manual analysis. The analytical outputs, including sentiment distributions, confusion matrices, and class-specific word clouds, provide interpretable information that facilitates policy evaluation and public opinion analysis. Furthermore, the proposed framework can serve as a foundation for future decision support systems or real-time social media monitoring platforms that continuously analyze public sentiment toward macroeconomic issues.

RESULT AND DISCUSSION

A. RESULT

The final corpus used in this study consists of 17,838 Indonesian-language YouTube comments collected from videos discussing macroeconomic policy issues, including inflation, exchange rate fluctuations, interest rate decisions, fiscal policy, and Bank Indonesia's monetary policy. Following the Data Preparation phase of the CRISP-DM methodology, the raw comments underwent a sequence of preprocessing steps, including case folding, removal of non-alphabetic characters, tokenization, Indonesian stopword removal, and stemming using the Sastrawi library to normalize inflected words into their root forms. Each comment was manually annotated into one of three sentiment categories: positive, neutral, or negative.

Subsequently, the preprocessed text was transformed into numerical representations using Term Frequency-Inverse Document Frequency (TF-IDF) vectorization with a maximum vocabulary size of 5,000 features, resulting in a feature matrix of 17,838 × 5,000. The dataset was then partitioned into training and testing subsets using an 80:20 stratified split, producing 14,270 training instances and 3,568 testing instances while preserving the original sentiment class distribution in both subsets.

As presented in Table 1, the dataset is characterized by an imbalanced class distribution. The neutral class constitutes the majority of the corpus with 11,070 comments (62.06%), followed by the positive class with 3,632 comments (20.36%) and the negative class with 3,136 comments (17.58%). The same distribution is maintained in the testing dataset through stratified sampling. This imbalance suggests that YouTube users are more likely to express descriptive, informational, or neutral opinions regarding macroeconomic policies than strongly polarized positive or negative sentiments. Consequently, model performance was evaluated using not only Accuracy but also Precision, Recall, and F1-score to provide a more comprehensive assessment under imbalanced class conditions.

Tabel 1. Sentiment class distribution in the initial dataset (before train-test split)

Sentiment Class	Number of Comments	Percentage
Neutral	11,070	62.06%
Positive	3,632	20.36%
Negative	3,136	17.58%
Total	**17,838**	**100%**

The XGBoost classifier was trained on the TF-IDF feature representation using a multi-class softmax objective. On the held-out test set of 3,568 comments, the model achieved an overall accuracy of 85.68%. Table 2 presents the complete classification report, including precision, recall, and F1-score for each sentiment class.

Tabel 2. Classification performance of the XGBoost model on the test set

Class	Precision	Recall	F1-Score	Support
Negative (0)	0.79	0.75	0.77	627
Neutral (1)	0.87	0.94	0.90	2214
Positive (2)	0.86	0.70	0.77	727
Accuracy			**0.86**	3568
Macro Avg	0.84	0.80	0.82	3568
Weighted Avg	0.86	0.86	0.85	3568

The model performed best on the neutral class, achieving a precision of 0.87, recall of 0.94, and F1-score of 0.90. This result is expected given that the neutral class was the majority class in the training data, providing the model with substantially more examples from which to learn its distinguishing features. The positive class achieved a comparatively high precision of 0.86 but a lower recall of 0.70, indicating that while the model rarely mislabels other classes as positive, it fails to identify a notable proportion of genuinely positive comments. The negative class showed a more balanced but slightly lower performance, with a precision of 0.79 and a recall of 0.75. The macro-average F1-score of 0.82 confirms that the model performs reasonably well across all three classes even when class sizes are not weighted, while the weighted-average F1-score of 0.85 reflects strong aggregate performance driven by the dominant neutral class.

Figure 2 and Table 3 present the confusion matrix generated from the model's predictions on the test set, which offers a more granular view of the classification errors than the aggregate metrics alone.

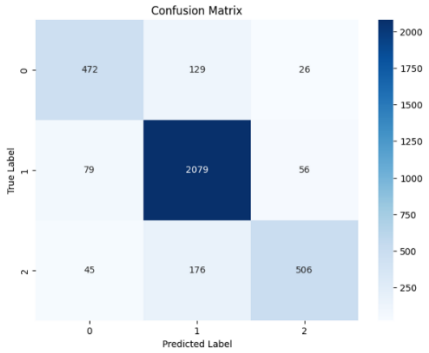


Figure 2. Confusion matrix of the XGBoost sentiment classification model

Tabel 3. Confusion matrix values (rows = true label, columns = predicted label)

True \ Predicted	Negative	Neutral	Positive
Negative	**472**	129	26
Neutral	79	**2079**	56
Positive	45	176	**506**

Of the 627 negative comments in the test set, 472 (75.3%) were correctly classified, while 129 (20.6%) were misclassified as neutral and 26 (4.1%) as positive. The neutral class showed the highest classification accuracy, with 2,079 of 2,214 comments (93.9%) correctly identified; the remaining cases were split between misclassification as negative (79 comments, 3.6%) and positive (56 comments, 2.5%). For the positive class, 506 of 727 comments (69.6%) were correctly classified, while a substantial share, 176 comments (24.2%), were misclassified as neutral, and 45 (6.2%) as negative.

This error pattern reveals that the dominant source of confusion in the model lies between the neutral and the positive/negative classes, rather than between positive and negative directly. In particular, the relatively high number of positive comments predicted as neutral (176 cases) is consistent with the recall gap observed for the positive class in Section 4.2 and suggests that many mildly positive or optimistic remarks about the economy (e.g., cautious approval of a policy) share vocabulary with neutral, descriptive comments. Confusion between the negative and positive classes was minimal (26 and 45 cases respectively), indicating that the model rarely confuses opposing sentiment polarities and that most of its errors are boundary cases along the neutral axis.

To complement the quantitative evaluation, word clouds were generated for each sentiment class from the preprocessed comment text Figures 3 to identify the most frequent terms associated with each category.



Figure 3. Wordcloud Sentiment Positive, Negative and Neutral

As illustrated in Figure 3, the three word clouds reveal distinct linguistic characteristics across the positive, neutral, and negative sentiment classes while sharing a common macroeconomic vocabulary. Across all sentiment categories, terms such as rupiah, Indonesia, negara (country), ekonomi (economy), pemerintah (government), and dollar appear prominently, indicating that discussions consistently revolve around Indonesia's macroeconomic conditions.

The positive sentiment word cloud is characterized by evaluative terms such as bagus (good), kuat (strong), percaya (trust), solusi (solution), and menguat (strengthening). These words suggest that positive comments primarily express optimism regarding economic stability, confidence in government policies, and expectations of a stronger national currency. The frequent occurrence of these positive expressions indicates supportive public attitudes toward macroeconomic policy and economic performance.

In contrast, the negative sentiment word cloud contains terms associated with economic deterioration and public dissatisfaction, including melemah (weakening), krisis (crisis), susah (difficult), miskin (poor), korupsi (corruption), and anjlok (plummeting). These expressions reflect concerns about currency depreciation, rising economic pressure, and dissatisfaction with government performance. The presence of informal and emotionally charged language further indicates that YouTube users often express criticism in a spontaneous and unfiltered manner.

The neutral sentiment word cloud is dominated by descriptive and topic-related terms, including presiden (president), Prabowo, Jokowi, ekspor (export), and conversational expressions such as gak, ya, nya, and aja. Unlike the positive and negative classes, these comments generally describe economic events, discuss political figures, or exchange information without conveying explicit positive or negative opinions.

An important observation from Figure 3 is the substantial lexical overlap among the three sentiment classes. High-frequency domain-specific terms, such as rupiah, Indonesia, ekonomi, and pemerintah, occur consistently across all categories because they represent the primary discussion topics rather than sentiment-bearing expressions. Consequently, the classification model relies on contextual and evaluative words to distinguish sentiment polarity. Although TF-IDF effectively emphasizes discriminative terms, its bag-of-words representation does not capture contextual semantics, making it challenging to differentiate comments that share similar topical vocabulary but express different sentiment orientations. This observation is consistent with the confusion matrix results, where the neutral class exhibits the greatest overlap with both the positive and negative classes.

B. DISCUSSION

Overall, the experimental results demonstrate that XGBoost combined with TF-IDF features provides an effective approach for sentiment classification of Indonesian YouTube comments related to macroeconomic policy. The proposed model achieved an overall accuracy of 85.68% and a weighted F1-score of 0.85, indicating robust classification performance despite the imbalanced distribution of sentiment classes. These results are consistent with previous studies on Indonesian social media sentiment analysis, where conventional machine learning algorithms combined with TF-IDF representations generally achieve competitive performance on noisy and informal textual data. The findings suggest that XGBoost effectively exploits sparse, high-dimensional feature spaces

by iteratively combining decision trees to learn discriminative lexical patterns associated with different sentiment categories.

A detailed analysis of the classification results reveals that most prediction errors occurred between the neutral class and the two polarized classes rather than between positive and negative sentiments. This behavior can be explained by several factors. First, the substantial class imbalance, where neutral comments account for more than 62% of the dataset, naturally biases the classifier toward the majority class, resulting in a high recall for neutral comments (0.94) while reducing the recall of the minority classes, particularly the positive class (0.70). Second, the word cloud analysis demonstrates considerable lexical overlap across sentiment categories. Domain-specific terms such as rupiah, Indonesia, government, and economy appear frequently in all three classes because they describe the discussion topic rather than the underlying sentiment. Consequently, TF-IDF representations alone cannot always distinguish whether these shared terms are accompanied by positive, negative, or neutral expressions.

Furthermore, the linguistic characteristics of Indonesian YouTube comments introduce additional challenges for sentiment classification. Many comments contain informal language, abbreviations, regional expressions, code-mixing, and occasional sarcasm, all of which increase semantic ambiguity. Although preprocessing techniques such as stemming and stopword removal reduce lexical variation, the bag-of-words nature of TF-IDF does not preserve contextual information or word order. As a result, comments expressing subtle approval or criticism are frequently represented by feature vectors that closely resemble neutral statements, explaining the concentration of misclassification along the neutral decision boundary observed in the confusion matrix.

From an application perspective, the proposed model demonstrates strong potential for supporting large-scale public opinion monitoring on macroeconomic policy. The achieved performance indicates that the model is sufficiently reliable for identifying overall sentiment distributions, tracking changes in public opinion following monetary or fiscal policy announcements, and assisting policymakers in understanding public responses through automated analysis of social media data. However, the relatively lower recall for the positive class suggests that mildly supportive comments are more likely to be classified as neutral. Therefore, applications requiring highly accurate identification of positive public reactions may benefit from incorporating manual verification for borderline cases.

Despite these promising results, several opportunities remain for improving classification performance. Future research may explore contextual language models such as IndoBERT or RoBERTa, which capture semantic relationships beyond individual word frequencies and may reduce confusion between neutral and polarized comments. Additionally, techniques for addressing class imbalance, including class-weighted learning or synthetic oversampling methods, could improve minority-class recognition. Expanding annotation guidelines to better capture sarcasm, implicit sentiment, and informal Indonesian expressions would also contribute to developing more robust sentiment classification models for social media discourse on macroeconomic policy.

CONCLUSION

This study developed and evaluated an XGBoost-based sentiment classification model for analyzing public opinion on Indonesian macroeconomic policy using 17,838 Indonesian-language YouTube comments. Experimental results demonstrate that the

proposed approach effectively classifies comments into positive, neutral, and negative sentiment categories, achieving an overall accuracy of 85.68% and a weighted F1-score of 0.85. The model showed the best performance on the neutral class, while confusion matrix and word cloud analyses revealed that most classification errors occurred between neutral and polarized comments because of shared macroeconomic vocabulary and subtle linguistic differences. These findings indicate that the combination of Indonesian-language text preprocessing, TF-IDF feature extraction, and XGBoost provides a reliable and scalable approach for large-scale sentiment analysis of macroeconomic discussions on social media.

Despite its promising performance, this study has several limitations. The dataset is characterized by an imbalanced sentiment distribution and relies on manually assigned sentiment labels, while the TF-IDF representation is unable to capture contextual semantics, sarcasm, and informal language commonly found in YouTube comments. Future research may improve classification performance by incorporating contextual language models such as IndoBERT or RoBERTa, applying class-balancing techniques, and expanding the annotation guidelines to better represent implicit sentiment and slang expressions. Furthermore, the proposed framework can serve as a foundation for real-time public opinion monitoring systems that support evidence-based evaluation of macroeconomic policies.

REFERENCES

Adilla, R. E., Huda, M., Aziz, M., & Suadaa, L. H. (2024). Aspect-Based Sentiment Analysis of Transportation Electrification Opinions on YouTube Comment Data. *Jurnal Aplikasi Statistika & Komputasi Statistik*, *16*(2). https://doi.org/10.34123/jurnalasks.v16i2.790

Alnofeli, K. K., Akter, S., Yanamandram, V., & Hani, U. (2026). AI-powered CRM capability model: Advancing marketing ambidexterity, profitability and competitive performance. *International Journal of Information Management*, *86*. https://doi.org/10.1016/j.ijinfomgt.2025.102981

Angdresey, A., Sitanayah, L., & Tangka, I. L. H. (2025). Sentiment Analysis for Political Debates on YouTube Comments using BERT Labeling, Random Oversampling, and Multinomial Naïve Bayes. *Journal of Computing Theories and Applications*, *2*(3). https://doi.org/10.62411/jcta.11668

Guo, K., Shen, C., Zhou, X., Ren, S., Hu, M., Shen, M., Chen, X., & Guo, H. (2024). Traffic Data-Empowered XGBoost-LSTM Framework for Infectious Disease Prediction. *IEEE Transactions on Intelligent Transportation Systems*, *25*(1). https://doi.org/10.1109/TITS.2022.3172206

Hadi Arfian, M., Sofu Maruhawa, H., Aisyah Putri, A., Ratulugina, D., Ridzky Khoirunnisa, E., Hermawan, V., & Julaibib Syahr, A. (2025). Analisis Sentimen Pada Media Sosial Menggunakan Metode Support Vector Machine. *Jurnal Ilmu Teknik Dan Komputer*, *09*(01).

Hama Aziz, R. H., & Dimililer, N. (2021). SentiXGboost: enhanced sentiment analysis in social media posts with ensemble XGBoost classifier. *Journal of the Chinese Institute of Engineers, Transactions of the Chinese Institute of Engineers,Series A*, *44*(6). https://doi.org/10.1080/02533839.2021.1933598

Kemp, S. (2025). *Digital 2025: global advertising trends*. Datareportal.

Le Dong Xuan, M. N., Thi Ngoc Bich, M. N., & Bao Phat, T. V. B. P. (2026). MEASURING THE IMPACT OF NON-TARIFF MEASURES ON VIETNAM'S AGRICULTURAL AND FRUIT EXPORTS IN THE CONTEXT OF

IMPLEMENTING NEW GENERATION FTAS: APPLYING THE GRAVITY MODEL. *International Journal of Leadership and Management*. https://doi.org/10.65176/ijlm.v2i2.01

Le, H. S., Do, T. V. H., Nguyen, M. H., Tran, H. A., Pham, T. T. T., Nguyen, N. T., & Nguyen, V. H. (2024). Predictive model for customer satisfaction analytics in E-commerce sector using machine learning and deep learning. *International Journal of Information Management Data Insights*, *4*(2). https://doi.org/10.1016/j.jjimei.2024.100295

Ma'aly, A. N., Pramesti, D., Fathurahman, A. D., & Fakhrurroja, H. (2024). Exploring Sentiment Analysis for the Indonesian Presidential Election Through Online Reviews Using Multi-Label Classification with a Deep Learning Algorithm. *Information (Switzerland)*, *15*(11). https://doi.org/10.3390/info15110705

Mandhasiya, D. G., Murfi, H., & Bustamam, A. (2024). The hybrid of BERT and deep learning models for Indonesian sentiment analysis. *Indonesian Journal of Electrical Engineering and Computer Science*, *33*(1). https://doi.org/10.11591/ijeecs.v33.i1.pp591-602

Moghadamnejad, A., Moghaddasi, M. A., Hamidia, M., Mohammadi, R. K., & Zare, M. (2026). Ranking Earthquake Prediction Algorithms: A Comprehensive Review of Machine Learning and Deep Learning Methods. In *Soil Dynamics and Earthquake Engineering* (Vol. 200). https://doi.org/10.1016/j.soildyn.2025.109740

Mohapatra, S., Mukherjee, R., Roy, A., Sengupta, A., & Puniyani, A. (2022). Can Ensemble Machine Learning Methods Predict Stock Returns for Indian Banks Using Technical Indicators? *Journal of Risk and Financial Management*, *15*(8). https://doi.org/10.3390/jrfm15080350

Mustofa, A., & Novita, R. (2022). Klasifikasi Sentimen Masyarakat Terhadap Pemberlakuan Pembatasan Kegiatan Masyarakat Menggunakan Text Mining Pada Twitter. *Building of Informatics, Technology and Science (BITS)*, *4*(1). https://doi.org/10.47065/bits.v4i1.1628

Pardede, E. T., Tania, K. D., & Afrina, M. (2026). Knowledge Discovery: Analisis Sentimen dan Emosi WhatsApp Business dengan Machine learning dan Deep Learning. *Jurnal Nasional Teknologi Dan Sistem Informasi*, *11*(3). https://doi.org/10.25077/teknosi.v11i3.2025.310-318

Patil, R., Boit, S., Gudivada, V., & Nandigam, J. (2023). A Survey of Text Representation and Embedding Techniques in NLP. *IEEE Access*, *11*. https://doi.org/10.1109/ACCESS.2023.3266377

Pratiwi, N., Hartanti, D., & Arum Sari, A. (2025). Sistem Rekomendasi Pemilihan Laptop Menggunakan Metode Content-Based Filtering Berbasis Spesifikasi Produk. *Infotek: Jurnal Informatika Dan Teknologi*, *8*(2). https://doi.org/10.29408/jit.v8i2.30618

sabbah Naiarni panjaitana, Azzikrarairyahannum, Syahilatuzzahira, P. I. (2024). Sentimen Analisis Pengguna Media Sosial Berdasarkan Metode Ekstraksi Fitur dan Klasifikasi. *Jurnal Ilmu Komputer*, *16*(2).

Samih, A., Ghadi, A., & Fennan, A. (2023). Enhanced sentiment analysis based on improved word embeddings and XGboost. *International Journal of Electrical and Computer Engineering*, *13*(2). https://doi.org/10.11591/ijece.v13i2.pp1827-1836

Sari, A. A., & Sari, O. K. (2026). *Pendekatan Interpretable Machine Learning untuk Analisis Keberhasilan Kampanye Pemasaran Menggunakan CatBoost dan SHAP*. *9*(2). https://doi.org/https://doi.org/10.29408/jit.v9i2.34715

Sari, O. K., & Cherid, A. (2023). Aplikasi Berbasis Web Menggunakan Apriori Untuk Rekomendasi Bundling Produk Sembako. *Bit (Fakultas Teknologi Informasi Universitas Budi Luhur)*, *20*(2). https://doi.org/10.36080/bit.v20i2.2444

Setyawan, Y. A., Nastiti, F. E., & Sari, A. A. (2025). Analisis Sentimen Ulasan Tempat Wisata Umbul Sigedang Di Google Maps Menggunakan Algoritma Naive Bayes. *Smart Comp :Jurnalnya Orang Pintar Komputer*, *14*(2).

Shivsharan, N., Kambli, V., Dabholkar, S., Dalvi, A., & Sukali, T. (2025). Evaluating Machine Learning Models for Sentiment Analysis of YouTube Comments and Creating an Accessible Web Application for Comment Analysis. *Procedia Computer Science*, *258*. https://doi.org/10.1016/j.procs.2025.04.574

Tangkudung, A. G. D., Rante, J. Z., Pakpahan, M., Putri, A., Puteri, A., Pribadi, D., Gosita, D. R., Rustandi, J., CPR, J. J., Ghalib, K., Farieanna, R., Purnawirawati, S., Setiawan, S., & Bassuni, W. (2024). INOVASI BISNIS PARIWISATA INDONESIA PASCA PANDEMI COVID-19. *Dinamika: Jurnal Manajemen Sosial Ekonomi*, *4*(1). https://doi.org/10.51903/dinamika.v4i1.433

Tsiourlini, M., Tzafilkou, K., Karapiperis, D., & Tjortjis, C. (2024). Text Analytics on YouTube Comments for Food Products. *Information (Switzerland)*, *15*(10). https://doi.org/10.3390/info15100599