TELKOMNIKA Telecommunication Computing Electronics and Control
Vol. 24, No. 3, June 2026, pp. 883~890
ISSN: 1693-6930, DOI: 10.12928/TELKOMNIKA.v24i3.27473

TikTok store affiliate performance sentiment analysis using support vector machine and gradient boosting machine methods

Fersellia[1], Fahmi Fachri[1], Afdhal Fauzan[1], Nihayatus Zaen[2]
[1]Department of Electrical Engineering, Ma'arif Nahdlatul Ulama University, Kebumen, Indonesia
[2]Department of History, Faculty of Cultural Sciences, Gadjah Mada University, Yogyakarta, Indonesia

Article Info

Article history:

Received Aug 20, 2025
Revised Nov 28, 2025
Accepted Dec 8, 2025

Keywords:

Affiliate marketing
Gradient booster machine
Sentiment analysis
Support vector machine
TikTok store

ABSTRACT

The development of social media-based e-commerce, particularly, opens new opportunities for digital affiliate systems. This study examines public perception of affiliate performance through comment sentiment analysis (positive, negative, neutral) using support vector machine (SVM) and gradient boosting machine (GBM). Data was collected from TikTok Shop comments, processed through text preprocessing, manual labeling, and then analyzed using Python. Evaluation using accuracy, precision, recall, and F1-score metrics showed that the combination of the synthetic minority oversampling technique (SMOTE) with SVM and GBM improved classification performance, although negative sentiment remained challenging. SVM achieved the highest accuracy (84%) with a ratio of 90:10, while GBM excelled in detecting neutral sentiment (F1 0.91). These findings are useful for sentiment-based marketing strategies and natural language processing (NLP) development for Indonesian-language texts on TikTok Shop.

Corresponding Author:

Fersellia
Department of Electrical Engineering, Ma'arif Nahdlatul Ulama University
Kutoarjo Street Km. 05, Wonoboyo, Jatisari, Kebumen Regency 54317, Central Java, Indonesia
Email: fersellia98@gmail.com

1. INTRODUCTION

The rapid advancement of digital technology has transformed marketing strategies, with social media emerging as a dominant platform for sales and affiliate promotion. TikTok Shop represents one of the fastest-growing examples of this phenomenon, enabling users to act as affiliates who earn commissions based on product sales through interactive short videos. This model effectively expands market reach, particularly among younger audiences, due to its algorithmic engagement and visual content dynamics [1].

However, affiliate performance on TikTok Shop is not solely determined by content frequency or follower count but is strongly influenced by public sentiment toward products, affiliates, and brands. Negative sentiment expressed through user comments can diminish consumer trust and reduce conversion rates [2]. Despite the growing relevance of social commerce, empirical research that systematically analyzes the relationship between user sentiment and affiliate performance on TikTok Shop remains scarce. Furthermore, existing studies tend to rely on general sentiment detection frameworks or focus on Western-language datasets, which are not fully applicable to the linguistic and cultural nuances of Indonesian social media users. These shortcomings emphasize the importance of having a locally relevant sentiment analysis model that can manage informal language, a lot of slang terms, and the context present in responses on TikTok Shop.

Important issues arise from not properly checking how well machine learning methods compare in this field. Even though support vector machines (SVM) and gradient boosting machines (GBM) have shown strong results when sorting text, their relative effectiveness in evaluating sentiments in Indonesian on TikTok Shop has yet to be validated through empirical research. As a result, pinpointing which model more effectively interprets user sentiments particularly when paired with advanced preprocessing and embedding methods represents a critical avenue for ongoing scholarly inquiry.

As a supervised learning algorithm, SVM operates by determining the most optimal hyperplane to achieve clear category separation. It works well in tasks where the data has many features, such as text classification [3], [4]. On the other hand, GBM uses a method that combines multiple models to gradually reduce errors in predictions, which makes it effective for dealing with complicated data sets [5], [6]. To make sure the data is of high quality, various steps are taken before training the models. These steps include breaking down text into words, removing unnecessary words, adjusting slang to standard language, and shortening words to their root forms, all of which help improve the accuracy of the models [7]-[10].

Sentiment analysis in this study was conducted using the valence aware dictionary and sentiment reasoner (VADER) approach. This instrument excels in interpreting the informal vernacular prevalent in social media content [11]-[13]. After the sentiment labels are assigned, the data is processed into numerical form using the term frequency–inverse document frequency (TF-IDF) technique. This technique quantifies word significance within the corpus [14], [15]. When an imbalance occurs because one category is overrepresented, the data is balanced by applying the synthetic minority oversampling technique (SMOTE). It generates synthetic instances of minority classes, thereby enhancing the model's overall efficacy [16]-[18].

Additionally, IndoBERT embeddings are employed to enhance the comprehension of word meanings and contextual nuances within the Indonesian language. This approach provides more accurate representation of language than traditional methods [19]-[22]. By combining traditional machine learning techniques with modern transformer models, the analysis of sentiment on social media platforms becomes more reliable.

The research also explores the potential of embedding sentiment engines into cyber-physical systems (CPS), where real-time sentiment feedback from IoT-based retail environments can inform dynamic promotional strategies and automated decision-making. Hence, this study not only contributes to computational sentiment classification but also highlights its practical implications in smart e-commerce systems. The reliability of the model is tested through various measures in the confusion matrix, namely accuracy, precision, recall, and F1-score, so that the evaluation results are more comprehensive [23], [24].

2. METHOD

The present research employs an experimental quantitative approach that integrates natural language processing (NLP) and machine learning techniques to categorize the sentiment of comments posted by TikTok shop users. The goal is to understand what people think about the TikTok shop affiliate program and to check how well different models work with various ways of representing text. Figure 1 shows the complete research design, illustrating the workflow from data collection through preprocessing, feature extraction, model training, optimization, evaluation, and deployment to the web dashboard.

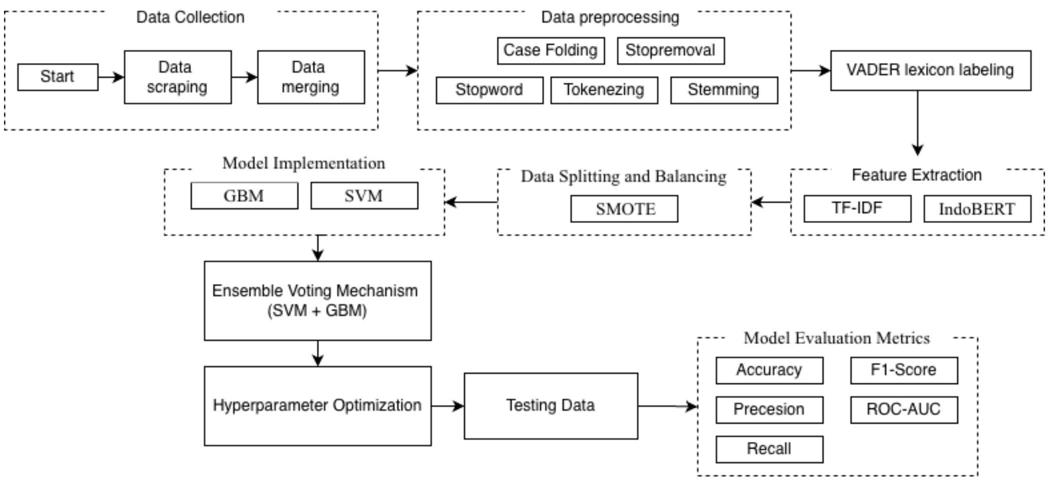

Figure 1. Research stages

a. Data collection: we gathered 9,208 comments from TikTok posts that had tags like "#tiktokshopaffiliate, #affiliatejualan", and similar hashtags using web scraping between January and July 2025. The comments came from both regular users and active affiliates, and we carefully selected them to make sure they were relevant to the affiliate program.
b. Data pre-processing: the gathered comments underwent preprocessing, which involved cleaning the data, segmenting them into individual words, stemming them to their root forms, and eliminating common words that contribute little to the overall meaning. Following this, each comment was assigned a sentiment classification as either positive, neutral, or negative.
c. Feature extraction: two different methods were used to represent the text: i) TF-IDF, which assigns numerical weights based on word frequency; and ii) IndoBERT embedding, which provides a more detailed and meaningful representation of Indonesian text by considering the context.
d. Data splitting and balancing: the dataset was divided into 80% for training and 20% for testing. The SMOTE technique was applied to balance the classes, while k-fold cross-validation was used to ensure stable and reliable evaluation.
e. Model training: the study employed two primary models: SVM and GBM. SVM excels in handling extensive and intricate datasets, whereas GBM is adept at identifying non-linear relationships [25]. Additionally, an ensemble model integrating SVM and GBM was evaluated, along with a transformer-based approach utilizing IndoBERT. The model is optimized using grid search and random search techniques to find the most effective configuration.
f. Model performance evaluation: performance evaluation was conducted using several metrics-accuracy, precision, recall, F1-score, and ROC–AUC. The research findings revealed that SVM produced the best overall accuracy, while GBM demonstrated the strongest ability to classify neutral sentiment. Meanwhile, the ensemble model produced consistent results across all sentiment classes.
g. Results integration: the final model was put into a web-based dashboard that helps small and medium businesses and affiliates track sentiment trends and assess their marketing strategies in real time.

3. RESULTS AND DISCUSSION

This study assesses the effectiveness of sentiment classification in user comments on TikTok shop using two machine learning methods: SVM and GBM. The dataset is divided into several training and testing schemes, such as 80:20, 90:10, and 70:30. To address class imbalance-particularly the dominance of neutral comments-the SMOTE is applied. The results are analyzed through a comparison of TF-IDF and IndoBERT performance, an examination of model errors, feature importance visualizations, and an interpretation of public sentiment trends concerning the TikTok shop affiliate program. Figure 2 illustrates the review data labeling process.

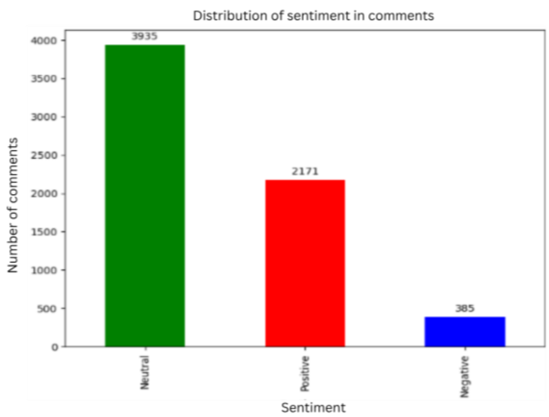

Figure 2. Review data labeling graph

3.1. Test results of SVM and GBM models with TF-IDF representation

The test results using TF-IDF, SMOTE, SVM, and GBM modeling show that the application of SMOTE is able to balance the class distribution that was originally dominated by neutral sentiment, so that the model can learn more fairly for all categories. Table 1 summarizes the detailed performance metrics (accuracy, F1-scores per class) across 80:20, 90:10, and 70:30 train-test splits for both models. In the 80:20

data proportion, SVM achieved 82% accuracy with the highest F1 score in the neutral class (0.86) and the lowest in the negative class (0.61), while GBM slightly outperformed with 83% accuracy and a higher neutral F1 (0.88), although performance in the negative class remained low (0.62). In the 90:10 scenario, SVM recorded the best performance with 84% accuracy and an increase in F1 in the positive class (0.80), while GBM remained relatively stable at 83% with a similar pattern, namely very good in the neutral class but weak in the negative class (0.59). In the 70:30 scenario, SVM accuracy decreased to 81% with a small improvement in the negative class (0.63), while GBM remained around 82% with strength in the neutral (0.87) and positive (0.76) classes, but again weakened in the negative class (0.60).

These differences in results are primarily due to the nature of each algorithm. SVM tends to be more consistent in maintaining performance balance between classes on high-dimensional data like TF-IDF, while GBM excels at capturing non-linear patterns, making it better for the majority (neutral) class. Although SMOTE effectively reduces bias toward the majority class, the diverse linguistic features and intricate semantic nature of negative comments hinder the model's ability to identify data structures or patterns, leading to inferior F1 scores for the negative class relative to others across all tested scenarios.

Table 1. Comparison of model test results (TF-IDF + SMOTE)

Proportion of training data: test	Model	Accuracy (%)	F1 positive	F1 negative	F1 neutral	Key notes
80:20	SVM	82	0.74	0.61	0.86	Good on neutral class, weak on negative
	GBM	83	0.75	0.62	0.88	Slightly higher accuracy, still weak in the negatives
90:10	SVM	84	0.80	0.60	0.88	Highest accuracy; significant improvement in positive
	GBM	83	0.77	0.59	0.88	Stable, but negative not increasing
70:30	SVM	81	0.72	0.63	0.86	More balanced, although accuracy is down
	GBM	82	0.76	0.60	0.87	Consistent, but still low in the negative

3.2. Test results of SVM and GBM models with IndoBERT embedding-based representation

Testing with IndoBERT embedding-based representations shows that both SVM and GBM are able to capture the semantic context of Indonesian better than TF-IDF, but performance between classes remains uneven. Table 2 presents the detailed performance metrics (accuracy, F1-scores per class) across 80:20, 90:10, and 70:30 train-test splits for SVM and GBM using IndoBERT embeddings. In the 80:20 scenario, SVM achieved 79% accuracy with the highest F1 score in the neutral class (0.86) and the lowest in the negative class (0.46), while GBM slightly outperformed with 80% accuracy and improved F1 scores in the negative (0.55) and positive (0.78) classes. This confirms GBM's superiority in handling non-linear patterns, although challenges remain in the negative class.

Table 2. Comparison of IndoBERT model test results

Data proportion	Model	Accuracy (%)	F1 negative	F1 neutral	F1 positive	Key notes
80:20	SVM	79	0.46	0.86	0.75	High neutral, very low negative
	GBM	80	0.55	0.85	0.78	Better than SVM in negative & positive
90:10	SVM	77	0.47	0.84	0.73	Accuracy down, negative still low
	GBM	81	0.55	0.86	0.79	Stable, excels in neutral & positive
70:30	SVM	81	0.63	0.86	0.74	More balanced, negative increases
	GBM	82	0.60	0.87	0.76	Most stable in all classes

In the 90:10 scenario, SVM accuracy actually decreased to 77%, despite consistently showing strength in the neutral class (F1 0.84). Performance in the negative class remained low (0.47), indicating the model's difficulty in recognizing the more linguistically diverse negative language variations. Conversely, the GBM model showed greater consistency in its accuracy at 81%, paired with a more balanced spread of F1-scores across the categories: 0.86 for neutral, 0.79 for positive, and 0.55 for negative. The stability of GBM in this scenario demonstrates that IndoBERT's representation is more optimal when combined with a boosting model capable of capturing interactions between features.

In the 70:30 scenario, SVM accuracy improved again to 80%, with the highest F1 score in the neutral class (0.86) and a significant improvement in the negative class (0.76), although the score in the positive class was relatively low (0.49). On the other hand, GBM again excelled with 82% accuracy and a more balanced F1 distribution, particularly for neutral (0.87) and positive (0.76), although its weakness in the negative class (0.60) remained apparent.

Overall, the IndoBERT embedding provides an advantage in the positive class and maintains consistent performance in the neutral class, but still struggles with the more linguistically diverse negative comments. SVM tends to be more adaptive to large data distributions, while GBM is more stable in maintaining average accuracy. This difference confirms that while transformer-based representations improve semantic understanding, the quality of minority data and stylistic variations remain decisive factors in sentiment classification.

3.3. Key findings

The findings of new things (novelty) can be formulated in the following aspects:

- Systematic comparison of TF-IDF vs IndoBERT on TikTok shop: this study not only uses the classic TF-IDF method but also compares it directly with the modern representation of IndoBERT embedding on the social media-based e-commerce domain (TikTok Shop), which has so far been rarely studied.
- Multi-scenario evaluation of train–test data with SMOTE: this study presents an in-depth analysis of three data proportion scenarios (80:20, 90:10, 70:30), combined with the SMOTE balancing technique. The results show consistent differences: SVM is more adaptive to variations in data proportions, while GBM is more stable in maintaining accuracy across all scenarios.
- Negative classroom performance as a unique challenge: a key finding is that both TF-IDF and IndoBERT still struggle to classify negative comments, even with balancing SMOTE. This indicates the linguistic diversity and unique style of negative comments on TikTok (e.g., slang, irony, or sarcasm), which are not fully accommodated by embedding models or classical algorithms.
- Avantages of GBM on non-linear representation of IndoBERT: the study found that GBM was more consistent than SVM when using IndoBERT embedding, with a more even F1-score distribution, especially in the positive and neutral classes. This confirms the combination of transformer and boosting representations as a promising pathway for Indonesian-language sentiment classification.
- Practical implications for E-commerce analytics: this study positions the sentiment model not only as an academic test but also as a tool to support micro, small, and medium enterprises (MSMEs) and TikTok Shop affiliates through potential integration into a real-time analytics dashboard system. This is a significant contribution because it connects NLP techniques with digital business needs.

4. FUTURE RESEARCH

Based on the challenges identified in this study, particularly the lower classification performance in the negative sentiment class, as well as the relevance of implementing sentiment results in digital systems, the following are suggestions for further research focused on innovation algorithms:

- Introducing advanced algorithmic innovation: to address existing limitations and improve semantic understanding of complex social media language (such as slang and sarcasm), future research needs to go beyond the TF-IDF and SMOTE approaches by integrating more sophisticated methodologies:
  a. Transformer models for richer semantic understanding: replacing feature extraction with end-to-end fine-tuning techniques using social media-specific Indonesian language variants of the transformer model, such as IndoBERTweet or IndoRoBERTa. These models offer much richer contextual semantic representation capabilities than TF-IDF, which is crucial for handling slang, irony, and ambiguity in TikTok comments.
  b. Real-time sentiment streaming: integrating sentiment engines into data streaming architectures using platforms like Apache Kafka or lightweight protocols like message queuing telemetry transport (MQTT). This integration is crucial for realizing a real-time affiliate performance monitoring system, enabling MSMEs or brands to respond to market sentiment instantly.
  c. Hybrid model (rule-based and machine learning): a hybrid model leverages the predictive power of machine learning (SVM/GBM/transformer) and a rule-based system built from a lexicon of sarcastic slang and emoticons. This hybrid approach can explicitly force the sentiment label to be negative if the rule criteria are met, thereby increasing recall in the minority (negative) class more consistently.
- Exploring hardware-software integration CPS: sentiment analysis should be framed as a digital sensor component within a CPS to unify and optimize affiliate and sales performance.
  a. CPS framework: sentiment results (from cyberspace) act as data inputs that trigger control actions in the physical environment.
  b. Promotional display control: if real-time sentiment analysis indicates a significant increase in positive sentiment for a specific product, the CPS system can send digital signals to actuators to control promotional displays in smart retail kiosks, such as turning on special lights or replaying affiliate videos on digital screens.

c. Pricing adjustment algorithms: high negative sentiment detected with specific keywords (“expensive,” “unreasonable price”) can trigger the cyber system to automatically adjust its pricing algorithm (dynamic pricing) or trigger temporary discounts to maintain conversions.
d. Triggering alerts in smart environments: negative sentiment focused on product or delivery quality (e.g., “damaged goods,” “long delivery”) should trigger urgent alerts sent to the supply chain management or logistics system.

– Implications for intelligent and multimodal systems: furthermore, enhanced machine sentiment can be at the core of CPS. This model can be embedded in edge computing devices (e.g., Raspberry Pi) as part of electronic and control systems in smart retail. Furthermore, in the context of TikTok, future research could analyze multimodal analysis by combining text sentiment with audio and visual signals (facial expressions) from affiliate video content using signal delivery and embedded vision techniques.

5. CONCLUSION

This research assesses the sentiment analysis of user comments on TikTok Shop by contrasting two text representation methods—TF-IDF embedding and IndoBERT—paired with two algorithms, namely SVM and GBM. The use of the SMOTE technique provides significant results in correcting the imbalance in class distribution, enhancing the model’s equity in managing various sentiment types.

The experimental findings indicate that the conventional lexical approach (TF-IDF + SVM) holds up as a strong benchmark, achieving the top accuracy of 84% on a 90:10 data split. The TF-IDF + GBM configuration exhibited the most consistent results in identifying neutral sentiment, with an F1 score of 0.88. On the other hand, the IndoBERT + GBM pairing delivered a more equitable performance across sentiment categories, especially for positive and neutral classes, boasting an average accuracy of 82%. Nonetheless, the negative class continues to pose significant difficulties for all models, owing to the linguistic variety, irony, and sarcasm prevalent in user feedback.

These findings provide empirical evidence that integrating contextual embedding IndoBERT with ensemble learning GBM can improve the stability and accuracy of sentiment detection in informal and meaningful Indonesian contexts. Furthermore, this study highlights practical implications for affiliate marketers and e-commerce platforms, where sentiment analysis results can be leveraged to develop more effective engagement strategies, improve customer experience, and support real-time decision-making in smart retail systems.

The originality of this research resides in its methodical evaluation of conventional versus transformer-driven methods within the TikTok Shop environment, delivering a noteworthy advancement to the progress of NLP for the Indonesian language. Future research is expected to explore domain-specific transformer models such as IndoBERTweet or IndoRoBERTa, as well as incorporate multimodal sentiment analysis based on text, audio, and visuals to further advance sentiment intelligence in Indonesia’s growing social commerce ecosystem.

FUNDING INFORMATION

This research was funded by PDP (domestic funding) after successfully passing the research selection process under Grant Number: [002/LPPM/UMNU/KP/V/2025].

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Fersellia	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓			✓	✓
Fahmi Fachri	✓	✓			✓	✓		✓	✓	✓	✓			
Afdhal Fauzan	✓	✓	✓	✓	✓	✓			✓	✓	✓			
Nihayatus Zaen	✓	✓			✓	✓		✓	✓	✓			✓	

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

The authors declare that there are no conflicts of interest, either financial or non-financial, that could influence the results or interpretation of this study.

DATA AVAILABILITY

The data supporting the findings of this study is openly available on [kaggle.com under the dataset name TikTok Shop Affiliate Comments Dataset] at https://www.kaggle.com/datasets/fersellia/dataset-tiktok-shop-affiliate-comments. This research was conducted with Google Colab using the SVM and GBM methods which can be accessed through the following link https://bit.ly/AnalisisSentimenTiktokshopaffiliate.

REFERENCES

[1] M. Srinivas, G. Sucharitha, A. Matta, and P. Chatterjee, Eds., *Machine Learning Algorithms and Applications*. Wiley, 2021, doi: 10.1002/9781119769262.
[2] P. Bakhsh, M. Ismail, M. A. Khan, M. Ali, and R. A. Memon, "Optimisation of sentiment analysis for E-commerce," *VFAST Transactions on Software Engineering*, vol. 12, no. 3, pp. 243–262, Sep. 2024, doi: 10.21015/vtse.v12i3.1907.
[3] B. Ghaddar and J. Naoum-Sawaya, "High dimensional data classification and feature selection using support vector machines," *European Journal of Operational Research*, vol. 265, no. 3, pp. 993–1004, Mar. 2018, doi: 10.1016/j.ejor.2017.08.040.
[4] L. Bharadwaj, "Sentiment analysis in online product reviews: mining customer opinions for sentiment classification," *International Journal For Multidisciplinary Research*, vol. 5, no. 5, Sep. 2023, doi: 10.36948/ijfmr.2023.v05i05.6090.
[5] R. Sibindi, R. W. Mwangi, and A. G. Waititu, "A boosting ensemble learning based hybrid light gradient boosting machine and extreme gradient boosting model for predicting house prices," *Engineering Reports*, vol. 5, no. 4, p. e12599, Apr. 2023, doi: 10.1002/eng2.12599.
[6] U. Singh, M. Rizwan, M. Alaraj, and I. Alsaidan, "A machine learning-based gradient boosting regression approach for wind power production forecasting: a step towards smart grid environments," *Energies*, vol. 14, no. 16, p. 5196, Aug. 2021, doi: 10.3390/en14165196.
[7] N. Garg and K. Sharma, "Text pre-processing of multilingual for sentiment analysis based on social network data," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 12, no. 1, p. 776, Feb. 2022, doi: 10.11591/ijece.v12i1.pp776-784.
[8] M. A. Palomino and F. Aider, "Evaluating the effectiveness of text pre-processing in sentiment analysis," *Applied Sciences*, vol. 12, no. 17, p. 8765, Aug. 2022, doi: 10.3390/app12178765.
[9] P. Krishnamoorthy, M. Sathiyanarayanan, and H. P. Proença, "A novel and secured email classification and emotion detection using hybrid deep neural network," *International Journal of Cognitive Computing in Engineering*, vol. 5, pp. 44–57, 2024, doi: 10.1016/j.ijcce.2024.01.002.
[10] M. A. Nq, L. P. Manik, and D. Widiyatmoko, "Stemming Javanese: another adaptation of the Nazief-Adriani algorithm," in *2020 3rd International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)*, IEEE, Dec. 2020, pp. 627–631, doi: 10.1109/ISRITI51436.2020.9315420.
[11] K. L. Tan, C. P. Lee, and K. M. Lim, "A survey of sentiment analysis: approaches, datasets, and future research," *Applied Sciences*, vol. 13, no. 7, p. 4550, Apr. 2023, doi: 10.3390/app13074550.
[12] D. C. Youvan, "Understanding sentiment analysis with VADER: a comprehensive overview and application," *AI and Data Science Journal*, no. June, 2024.
[13] V. Nurcahyawati, Z. Mustaffa, and M. Khalaf, "Exceeding manual labeling: VADER lexicon as an accurate alternative to automatic sentiment classification," *The International Arab Journal of Information Technology*, vol. 22, no. 2, pp. 225–235, 2025, doi: 10.34028/iajit/22/2/2.
[14] Z. Jiang, B. Gao, Y. He, Y. Han, P. Doyle, and Q. Zhu, "Text classification using novel term weighting scheme-based improved TF-IDF for internet media reports," *Mathematical Problems in Engineering*, vol. 2021, pp. 1–30, Mar. 2021, doi: 10.1155/2021/6619088.
[15] S. Sarica and J. Luo, "Stopwords in technical language processing," *PLOS ONE*, vol. 16, no. 8, p. e0254937, Aug. 2021, doi: 10.1371/journal.pone.0254937.
[16] D. Elreedy, A. F. Atiya, and F. Kamalov, "A theoretical distribution analysis of synthetic minority oversampling technique (SMOTE) for imbalanced learning," *Machine Learning*, vol. 113, no. 7, pp. 4903–4923, Jul. 2024, doi: 10.1007/s10994-022-06296-4.
[17] S. Wang, Y. Dai, J. Shen, and J. Xuan, "Research on expansion and classification of imbalanced data based on SMOTE algorithm," *Scientific Reports*, vol. 11, no. 1, p. 24039, Dec. 2021, doi: 10.1038/s41598-021-03430-5.
[18] H. Hairani, T. Widiyaningtyas, and D. Dwi Prasetya, "Addressing class imbalance of health data: a systematic literature review on modified synthetic minority oversampling technique (SMOTE) strategies," *JOIV : International Journal on Informatics Visualization*, vol. 8, no. 3, pp. 1310–1318, Sep. 2024, doi: 10.62527/joiv.8.3.2283.
[19] J. Abramson *et al.*, "Accurate structure prediction of biomolecular interactions with AlphaFold 3," *Nature*, vol. 630, no. 8016, pp. 493–500, Jun. 2024, doi: 10.1038/s41586-024-07487-w.
[20] G. M. Foody, "Challenges in the real world use of classification accuracy metrics: from recall and precision to the Matthews correlation coefficient," *PLOS ONE*, vol. 18, no. 10, p. e0291908, Oct. 2023, doi: 10.1371/journal.pone.0291908.
[21] E. Casas, L. Ramos, E. Bendek, and F. Rivas-Echeverría, "Assessing the effectiveness of YOLO architectures for smoke and wildfire detection," *IEEE Access*, vol. 11, pp. 96554–96583, 2023, doi: 10.1109/ACCESS.2023.3312217.
[22] D. J. Hand, P. Christen, and N. Kirielle, "F*: an interpretable transformation of the F-measure," *Machine Learning*, vol. 110, no. 3, pp. 451–456, Mar. 2021, doi: 10.1007/s10994-021-05964-1.
[23] J. Li, "Area under the ROC Curve has the most consistent evaluation for binary classification," *PLOS ONE*, vol. 19, no. 12, p. e0316019, Dec. 2024, doi: 10.1371/journal.pone.0316019.
[24] T. Poisot, "Guidelines for the prediction of species interactions through binary classification," *Methods in Ecology and Evolution*, vol. 14, no. 5, pp. 1333–1345, May 2023, doi: 10.1111/2041-210X.14071.
[25] M. D. Algifahri and S. Sriani, "Performance Analysis of Support Vector Machine with Hyperparameter Tuning for Engagement Rate Classification of TikTok Digital Marketing Content," *Inform : Jurnal Ilmiah Bidang Teknologi Informasi Dan Komunikasi*, vol. 11, no. 1, pp. 111–120, Feb. 2026, doi: 10.25139/inform.v11i1.11362.

BIOGRAPHIES OF AUTHORS

Fersellia earned a Bachelor's degree in Information Systems from Muhammadiyah University of Purworejo in 2020, after studying since 2017. She then continued her education and earned a Master's degree in Computer Science (M.Kom) in Informatics Engineering from Amikom University Yogyakarta in 2023. She is currently a lecturer in the Informatics Engineering Study Program, Ma'arif Nahdlatul Ulama University Kebumen (UMNU Kebumen). Her research focus and teaching expertise include artificial intelligence, natural language processing, machine learning, and digital marketing. Her research interests include intelligent system development, AI applications in digital marketing, and text-based data analysis. She can be contacted at email: fersellia98.gmail.com.

Fahmi Fachri is a lecturer at Universitas Ma'arif Nahdlatul Ulama (UMNU) Kebumen, Central Java, Indonesia. He earned his Master's degree from Universitas Ahmad Dahlan. He can be contacted at email: fahmifachriumnu@gmail.com.

Afdhal Fauzan is an Informatics Engineering student at Universitas Ma'arif Nahdlatul Ulama (UMNU) Kebumen. He began his undergraduate studies in 2023, following his graduation from SMK Ma'arif 9 Kebumen in 2020, where he majored in Software Engineering (RPL). With expertise in website and program development, he actively contributes to the field of informatics. He can be contacted at email: aftsu14@gmail.com.

Nihayatus Zaen earned a Bachelor's degree in Humanities from UIN Sunan Kalijaga Yogyakarta in 2022. After graduating, she began her media career as a content creator and affiliate on the TikTok platform. Although not entirely relevant to her major, the academic world has equipped her with excellent public speaking skills, which has allowed her career as an affiliate to continue to grow over time. Until 2025, this job became her main source of income, which also led her to pursue Master's degree at Gadjah Mada University. In addition to content creation, she is currently a student in the Master of History Program, Faculty of Cultural Sciences, UGM. Her research interests focus on the study of female ulama and Islamic boarding schools. She can be contacted at email: zaynia16@gmail.com.