Brilliance
Research of Artificial Intelligence

E-ISSN : 2807-9035
Volume 6, **Number** 1, May 2026
https://doi.org/10.47709/brilliance.v6i1.8073

Sentiment Analysis of TikTok Netizens on Oil Palm Issues in Papua Using KNN

Annisa Karima[1*], Naufal Abdulillah[2], Aril Maulana[3], Satria Harry Menov[4], T. Sukma Achriadi Sukiman[5]
[1,2,3,4] Information System Departement, University of Malikussaleh, Indonesia, [5]Informtics Department, University of Malikussaleh, Indonesia
[1]annisakarima@unimal.ac.id

open access

***Corresponding Author**

Article History:
Submitted: 11-02-2026
Accepted: 22-03-2026
Published: 21-04-2026

Keywords:
Sentiment analysis; TikTok; TF-IDF; KNN; oil palm; Papua.

ABSTRACT

Oil palm plantation in Papua has become a controversial issue that has generated diverse responses from the public, particularly regarding concerns over environmental impacts and ecosystem sustainability. These differing perspectives create the need to comprehensively and objectively understand public perceptions. This study aims to analyze the sentiment of Indonesian netizens toward the policy of oil palm plantation in Papua using a machine learning–based sentiment analysis approach. The research data were collected from user comments on the TikTok platform, which were subsequently processed through preprocessing stages, translation into English, and automatic labeling using TextBlob. The labeled data were then represented using Term Frequency–Inverse Document Frequency (TF-IDF) weighting and classified using the K-Nearest Neighbors (KNN) algorithm. The classification results using the K-Nearest Neighbors (KNN) algorithm indicate that out of a total of 220 data samples, 31 data (14%) were classified as positive sentiment and 189 data (86%) as negative sentiment. The classification process using the K-Nearest Neighbors (KNN) algorithm with the optimal K value of 5 achieved an accuracy of 77.27%, with a precision of 74.90%, recall of 66.00%, and an F1-score of 67.65%. The recall value for the positive class is relatively low, at 38.46%, indicating that the model still faces challenges in correctly identifying all positive data. This limitation is attributed to the imbalance in data distribution and the complexity of language used in social media comments. Nevertheless, the overall classification results suggest that the majority of netizens tend to oppose oil palm plantation in Papua, mainly due to concerns about environmental impacts and ecosystem degradation.

INTRODUCTION

Papua is one of the regions in Indonesia that possesses extensive tropical forest cover with exceptionally high levels of biodiversity. The soil conditions in Papua are generally dominated by red-yellow podzolic soils and latosols, which exhibit moderate to low fertility, thereby requiring special management practices for large-scale agricultural activities(BAPPENAS et al., 2024). In recent years, there has been discourse on the development of oil palm plantations in Papua as part of efforts to promote economic growth and equitable national development. However, this policy has sparked debate among the public due to its potential environmental and social impacts.

Over the past several decades, Indonesia's forests have faced numerous severe challenges that threaten the stability of their ecosystems. Deforestation driven by land expansion for agriculture, oil palm plantations, mining activities, as well as infrastructure development and human settlements has resulted in massive forest cover loss. This condition has led to a decline in biodiversity, disruption of ecosystem functions, and increased carbon emissions into the atmosphere, thereby accelerating the pace of climate change (Rahmawati A. Damiti et al., 2025).

In Aceh Province, the conversion of forest land into oil palm plantations has contributed to an increased risk of flooding and landslides due to the reduced capacity of the soil to absorb rainfall. Geologically, the Sumatra region is predominantly characterized by alluvial and podzolic soils, which generally exhibit low infiltration capacity, making the area highly susceptible to flooding when natural vegetation cover is diminished. This indicates that large-scale oil palm cultivation in regions with specific soil characteristics can significantly increase the potential for hydrometeorological disasters, such as floods and landslides, thereby raising concerns should similar practices be implemented in Papua.

The policy discourse of the President of the Republic of Indonesia regarding the development of oil palm plantations in Papua has elicited diverse responses from the public, both in support of and in opposition to the initiative. Therefore, sentiment analysis is essential to obtain an objective overview of public perceptions toward this policy. Sentiment analysis is a computational method used to identify public opinions, attitudes, and emotions expressed in textual form on social media platforms (Purnamasari et al., 2023). This approach enables more comprehensive data driven decision making and enhances responsiveness to public aspirations. In this study, data were collected from the

E-ISSN : 2807-9035
Volume 6, **Number** 1, May 2026
https://doi.org/10.47709/brilliance.v6i1.8073

TikTok social media platform, as it is one of the fastest-growing social media platforms in Indonesia and exhibits a high level of user engagement in disseminating public opinions(Indri Anggraini & Prima Mulyasari Agustini, 2025). The sentiment classification process in this study was carried out using the K-Nearest Neighbors (KNN) method, as this algorithm offers advantages in terms of implementation simplicity, flexibility across various data types, and competitive accuracy in text classification tasks (Murtiwiyati et al., 2025). In addition, KNN is capable of classifying new data based on pattern proximity to the training data, making it effective for sentiment analysis of dynamic social media text.

This study aims to analyze public perceptions and sentiment tendencies toward the discourse on the development of oil palm plantations in Papua by utilizing comment data from the TikTok platform and employing the K-Nearest Neighbors (KNN) classification method, as well as to evaluate the model's performance in accurately classifying public opinions.

LITERATURE REVIEW

Sentiment Analysis

Sentiment analysis is a computational study that aims to identify, extract, and interpret individuals' opinions, emotions, and attitudes toward a particular topic based on user-generated textual data, especially from social media and other online platforms. (Li et al., 2023). This approach is employed to automatically and systematically understand public perceptions of social, economic, political, and environmental issues. The primary objective of sentiment analysis is to classify opinion polarity into positive, negative, or neutral categories, thereby providing a quantitative overview of public attitudes toward a particular event or policy. The sentiment analysis process generally involves several main stages, including data collection, text preprocessing, feature extraction, and classification using machine learning or deep learning methods, such as Naive Bayes, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and neural network–based models. The application of sentiment analysis offers various benefits, including assisting policymakers in formulating more responsive strategies, enhancing the understanding of public opinion, and serving as a basis for evaluating the social impacts of policies or public discourse, particularly in the context of environmental issues and sustainable development (Alvionika et al., 2024).

Text Mining

Text mining is a process of extracting useful information, patterns, and knowledge from unstructured textual data by employing computational, statistical, and artificial intelligence techniques. It aims to transform raw text data into structured information that can be further analyzed to support decision-making. In the context of sentiment analysis, text mining plays a crucial role in processing textual data derived from social media, user reviews, and online comments so that they can be classified into specific sentiment categories. This process involves a series of Natural Language Processing (NLP) stages, including data cleaning, feature extraction, and modeling using machine learning algorithms(Purnamasari et al., 2023).

Preprocessing

Preprocessing is a crucial initial stage in text mining, as it aims to clean and prepare textual data for subsequent analysis. This stage is conducted to reduce noise, improve data quality, and enhance the accuracy of classification models. Preprocessing generally includes several key steps, such as case folding, tokenization, stopword removal, and stemming. These processes are intended to transform raw text into a simpler, more consistent, and meaningful data representation, thereby improving the performance of machine learning algorithms employed in sentiment analysis(Purnamasari et al., 2023).

TextBlob-Based Labeling

Data labeling is a crucial stage in sentiment analysis research, aiming to assign sentiment classes to each textual data instance according to predefined categories, namely positive, negative, and neutral. This labeling process serves as the foundation for constructing a labeled dataset, which is subsequently utilized in the training and testing phases of the classification model. Labeling can be performed either manually or automatically. In this study, labeling was conducted automatically using a lexicon-based approach through the TextBlob library.

The lexicon-based approach operates by utilizing a sentiment lexicon that contains a list of words with positive and negative connotations. The polarity of a text is determined by calculating the sentiment scores of the words contained within it. If the number or cumulative score of positive words exceeds that of negative words, the text is classified as positive sentiment. Conversely, if the number or cumulative score of negative words is greater, the text is categorized as negative sentiment, while texts with balanced scores are classified as neutral.

TextBlob is a Python-based Natural Language Processing (NLP) library that provides automatic sentiment analysis by computing polarity values within a range of −1 to +1. These polarity values are used to determine the emotional tendency of a text, where positive values indicate positive sentiment, negative values indicate negative sentiment, and values close to zero indicate neutral sentiment. This approach facilitates the labeling process for large-scale datasets, thereby improving time efficiency and ensuring consistency in labeling results (Purnamasari et al., 2023).

E-ISSN : 2807-9035
Volume 6, Number 1, May 2026
https://doi.org/10.47709/brilliance.v6i1.8073

K-Nearest Neighbor (KNN) Algorithm

The K-Nearest Neighbor (KNN) algorithm is a supervised learning method used for data classification based on the proximity between test data and training data. The working principle of KNN is to determine the class of a new data instance according to the majority class among its K nearest neighbors. The classification process is carried out by calculating the distance between the test data and all training data, then selecting the K data points with the smallest distances as the basis for class assignment. This method is simple, does not require complex model training, and is capable of producing relatively accurate classification results, particularly for numerical data that have undergone preprocessing and normalization stages (Nisa et al., 2025).

METHOD

The following figure illustrates the research methodology flowchart, which describes the sequential stages conducted in this study, starting from data collection, preprocessing, sentiment labeling, feature weighting using TF-IDF, classification using the K-Nearest Neighbors (KNN) algorithm, and model evaluation.

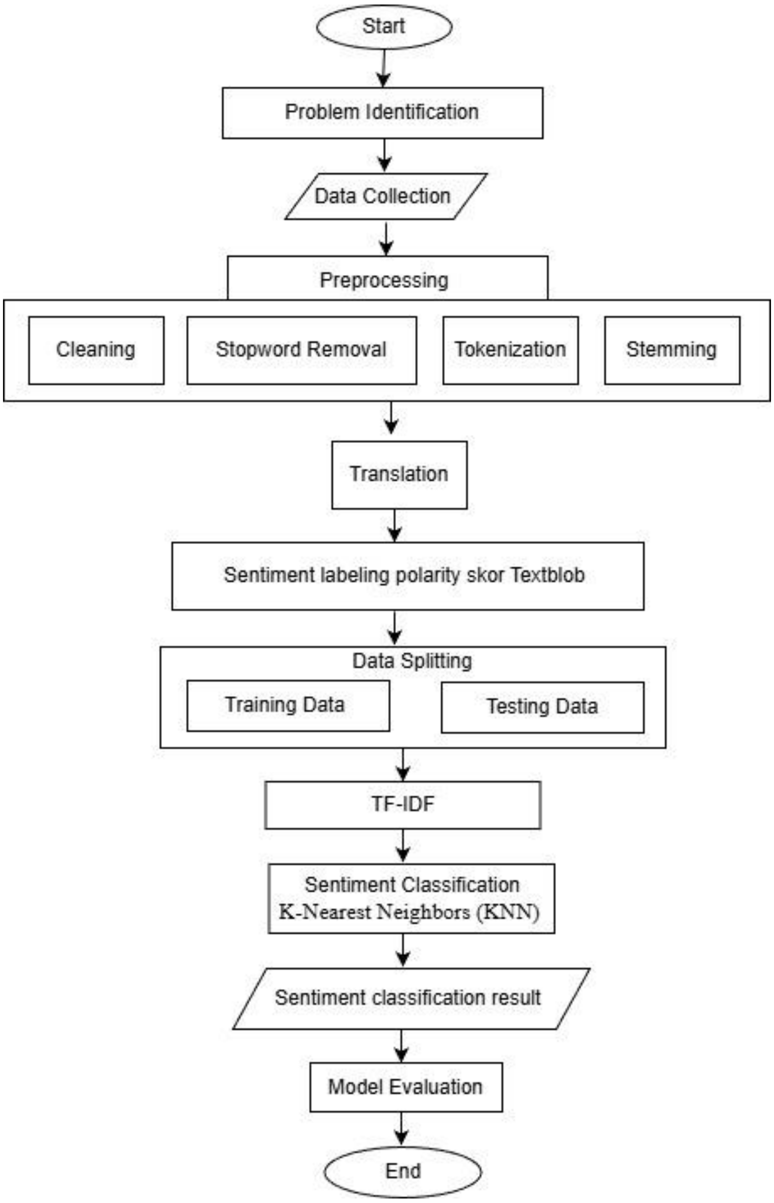

Figure 1. Research Methodology Flowchart

Problem Identification

The development of oil palm plantations in Papua has generated diverse public responses, ranging from support for potential economic growth and improved welfare to concerns over environmental impacts, ecosystem degradation, and threats to the sustainability of indigenous communities. The expansion of oil palm plantations is considered to have

E-ISSN : 2807-9035
Volume 6, **Number** 1, May 2026
https://doi.org/10.47709/brilliance.v6i1.8073

the potential to trigger deforestation, loss of biodiversity, and increased risks of environmental disasters, such as flooding and soil degradation. To understand public sentiment toward this policy, sentiment analysis based on social media data is required, as it reflects public opinion in real time. In this study, TikTok comment data are utilized as the primary source of information, while the K-Nearest Neighbors (KNN) method is employed as the classification technique due to its simplicity and effectiveness in grouping data based on textual pattern similarity.(Ayuk Puspita et al., 2025).

Data Collection

The data collection technique employed in this study was web scraping, with data sourced from the TikTok social media platform. Data retrieval was conducted using the search keyword *"penanaman kelapa sawit di Papua"* ("oil palm plantation in Papua"), resulting in a total of 220 collected comments. These data represent raw textual data; therefore, further preprocessing was required to ensure data quality and suitability for in-depth analysis.

Figure 2. Workflow of Data Collection

Pre-processing

Preprocessing is the initial stage in textual data processing, aimed at cleaning and preparing data for subsequent analysis. Social media data generally contain noise such as symbols, abbreviations, spelling errors, and non-standard words; therefore, data cleaning is required to improve data quality. The preprocessing stages include case folding, tokenization, stopword removal, and stemming, which are intended to produce more structured, concise, and relevant textual data. This stage has a significant impact on the performance of classification algorithms, as the quality of input data directly determines the accuracy level of sentiment analysis results (Hakim, 2021).

- **Data Cleaning**

The cleaning stage aims to remove irrelevant characters from the text, such as punctuation marks, numbers, URLs, mentions, hashtags, emojis, and other symbols (Randy Suryono, 2025). This process is essential to reduce noise that may interfere with the analysis and to improve data quality. Consequently, the resulting text contains only words that carry meaningful information for sentiment analysis (Karima et al., 2025).

Table 1. Data Cleaning

Data	Cleaning Results
SAWITTðŸ˜,TPI MINYK MAHALLL	sawitt tpi minyk mahalll
oh trrnyata yang rusak bukan alamnya, tapi presidennya ðŸ¤¦ðŸ□»â€□â™€ï¸□	oh trrnyata yang rusak bukan alamnya tapi presidennya
LANSIA INI SEDANG BERBISNIS BUKAN SEDANG MEMIMPIN NEGERI INI...	lansia ini sedang berbisnis bukan sedang memimpin negeri ini
kami rakyat Indonesia, menetapkan Prabowo Gibran sbg bencana nasional	kami rakyat indonesia menetapkan prabowo gibran sbg bencana nasional
udah kubilang dia cuma mikir uang bukan rakyat	udah kubilang dia cuma mikir uang bukan rakyat

- **Stopword removal**

The stopword removal stage aims to eliminate common words that do not have a significant influence on sentiment analysis, such as "dan," "yang," "di" "ke," and "dari." Removing stopwords enables the model to focus more on words that contain important information related to sentiment polarity, thereby improving the efficiency and performance of the classification model.

Table 2. Stopword removal result

Data Cleaning result	Stopword removal result
sawitt tpi minyk mahalll	sawitt tpi minyk mahalll
oh trrnyata yang rusak bukan alamnya tapi presidennya	trrnyata rusak bukan alamnya presidennya
lansia ini sedang berbisnis bukan sedang memimpin negeri ini	lansia sedang berbisnis bukan sedang memimpin negeri
kami rakyat indonesia menetapkan prabowo gibran sbg bencana nasional	rakyat indonesia menetapkan prabowo gibran sbg bencana nasional
udah kubilang dia cuma mikir uang bukan rakyat	udah kubilang cuma mikir uang bukan rakyat

E-ISSN : 2807-9035
Volume 6, **Number** 1, May 2026
https://doi.org/10.47709/brilliance.v6i1.8073

- **Tokenization**

The tokenization stage is the process of breaking text into individual word units or tokens. This process aims to facilitate the analysis of each word independently within a document. For instance, a sentence is divided into multiple tokens, each representing a single word. Tokenization enables more effective word frequency calculation and feature extraction.

Table 3. Tokenization

Stopword removal result	Tokenization
sawitt tpi minyk mahalll	[‘sawitt’, ‘tpi’, ‘minyk’, ‘mahalll’,]
trrnyata rusak bukan alamnya presidennya	[‘trrnyata’, ‘rusak’, ‘bukan’, ‘alamnya’, ‘presidennya’]
lansia sedang berbisnis bukan sedang memimpin negeri	[‘lansia’, ‘sedang’, ‘berbisnis’, ‘bukan’, ‘sedang’, ‘memimpin’, ‘negeri’]
rakyat indonesia menetapkan prabowo gibran sbg bencana nasional	[‘rakyat’, ‘indonesia’, ‘menetapkan’, ‘prabowo’, ‘gibran’, ‘sbg’, ‘bencana’, ‘nasional’]
udah kubilang cuma mikir uang bukan rakyat	[‘udah’, ‘kubilang’, ‘Cuma’, ‘mikir’, ‘uang’, ‘bukan’, ‘rakyat’]

- **Stemming**

The stemming stage aims to transform inflected or affixed words into their root forms. This process is performed to reduce word form variations and enhance data consistency(Ardyani Fitria Endah, 2022). For example, the words *“membantu”* (to help), *“dibantu”* (helped), and *“bantuan”* (assistance) are converted into the root word *“bantu”* (help). Consequently, the analysis and classification processes can be carried out more optimally.

Table 4.Stemming result

Tokenization	Stemming
[‘sawitt’, ‘tpi’, ‘minyk’, ‘mahalll’,]	sawitt tpi minyk mahalll
[‘trrnyata’, ‘rusak’, ‘bukan’, ‘alamnya’, ‘presidennya’]	trrnyata rusak bukan alam presiden
[‘lansia’, ‘sedang’, ‘berbisnis’, ‘bukan’, ‘sedang’, ‘memimpin’, ‘negeri’]	lansia sedang bisnis bukan sedang pimpin negeri
[‘rakyat’, ‘indonesia’, ‘menetapkan’, ‘prabowo’, ‘gibran’, ‘sbg’, ‘bencana’, ‘nasional’]	rakyat indonesia tetap prabowo gibran sbg bencana nasional
[‘udah’, ‘kubilang’, ‘Cuma’, ‘mikir’, ‘uang’, ‘bukan’, ‘rakyat’]	udah bilang cuma mikir uang bukan rakyat

Translation

The translation stage was performed by converting all preprocessed text into English using an API-based automatic translation service. This process aims to improve the quality of sentiment labeling using the TextBlob library, which is specifically developed and optimized for the English language. By performing translation, the sentiment labeling results are expected to be more accurate and consistent compared to directly processing Indonesian-language text.

Table 5. Translation result

Data	Translation
sawitt tpi minyk mahalll	Palm oil is expensive
trrnyata rusak bukan alam presiden	oh the real broken nature of the president
lansia sedang bisnis bukan sedang pimpin negeri	business seniors lead the country
rakyat indonesia tetap prabowo gibran sbg bencana nasional	The Indonesian people are still Prabowo responsible for the national disaster
udah bilang cuma mikir uang bukan rakyat	I already said that I think it's people's money

Sentiment labeling

Sentiment labeling was performed automatically using the TextBlob library, which generates polarity values ranging from −1 to +1. Data with polarity values greater than 0 were categorized as positive, values equal to 0 as neutral, and values less than 0 as negative. For the purpose of binary classification, the neutral class was merged into the negative class. This approach was adopted to simplify the classification process and enhance the clarity of sentiment polarity.

E-ISSN : 2807-9035
Volume 6, **Number** 1, May 2026
https://doi.org/10.47709/brilliance.v6i1.8073

Table 6. Labeling result

Translation	Klasifikasi Sentiment
Palm oil is expensive	negative
oh the real broken nature of the president	negative
business seniors lead the country	neutral
The Indonesian people are still Prabowo responsible for the national disaster	positive
I already said that I think it's people's money	neutral

Data Splitting

The dataset was divided into training and testing sets, consisting of 176 training samples and 44 testing samples. In sentiment analysis research, data splitting plays a crucial role in evaluating the generalization capability of the developed model. Generally, the dataset is divided into two parts, with 80% used as training data and 20% as testing data. The training data are utilized to build and train the KNN model so that it can classify comments into positive or negative sentiment categories. Meanwhile, the testing data are employed to evaluate the model's performance on previously unseen data, thereby minimizing the risk of overfitting. Through this data partitioning strategy, the model is expected not only to perform optimally on the training data but also to achieve high accuracy when processing new data.

TF-IDF

The Term Frequency–Inverse Document Frequency (TF-IDF) method was employed as a weighting technique for data that had undergone the preprocessing stage (Wong et al., 2025). The combination of TF-IDF and the K-Nearest Neighbor (KNN) algorithm has been proven to enhance performance in text classification tasks. In addition to its high effectiveness, TF-IDF provides a simple weighting mechanism while still delivering satisfactory classification performance (Sari et al., 2021).

Sentiment classification modeling

Sentiment classification modeling in this study was conducted using the K-Nearest Neighbors (KNN) algorithm to obtain an optimal classification model. The selection of the K parameter was carried out through an experimental process by testing several odd values of K, ranging from K = 1 to K = 21, using the macro F1-score as the evaluation metric. This metric was chosen to achieve balanced classification performance across all classes, particularly under conditions of imbalanced data distribution.

The experimental results indicate that K = 5 and K = 11 yielded the highest macro F1-score of 0.6765. However, in this study, K = 5 was selected as the optimal parameter value because it provided stable performance with lower model complexity. With K = 5, the KNN model achieved an accuracy of 77.27%, a precision of 0.7491, a recall of 0.6600, and an F1-score of 0.6765.

Based on the classification evaluation results, the KNN model demonstrated excellent performance in classifying negative sentiment, achieving a recall of 93.55% and an F1-score of 85.29%. Meanwhile, performance in the positive sentiment class remained relatively lower, with an F1-score of 50.00%, which was influenced by the imbalance in the number of data samples across classes. Overall, these results indicate that the KNN algorithm with K = 5 is capable of delivering sufficiently optimal classification performance in sentiment analysis.

Model Evaluation

Model evaluation is a crucial step in assessing the performance of the classification model based on the classification results. In this study, a confusion matrix was employed as the evaluation framework. The performance metrics used to measure the model include accuracy, precision, recall, and F1-score (Pinem & Putra, 2025). The following equations present the formulas used to evaluate the performance of the classification model, including the calculation of accuracy, precision, recall, and F1-score.

- **Accuracy**

Accuracy is used to measure the overall correctness of the model in classifying data, defined as the ratio of the number of correct predictions to the total number of tested data.

$$\text{Accuracy} = \frac{TP + TN}{TP+FP+TN+FN} \quad (1)$$

Information :

TP = True Positive (Actual positive class that is correctly predicted as positive).
TN = True Negative (Actual negative class that is correctly predicted as negative).
FP = False Positive (Actual negative class that is incorrectly predicted as positive).

E-ISSN : 2807-9035
Volume 6, **Number** 1, May 2026
https://doi.org/10.47709/brilliance.v6i1.8073

FN = False Negative (Actual positive class that is incorrectly predicted as negative).

- **Precision**

Precision is used to indicate how accurately the model predicts data belonging to the positive class, defined as the proportion of correct positive predictions to the total number of positive predictions generated by the model.

$$\text{Precision} = \frac{TP}{TP+FP} \tag{2}$$

Information :
TP = True Positive (TP) (Actual positive class that is correctly predicted as positive).
FP = False Positive (FP) (Actual negative class that is incorrectly predicted as positive).

- **Recall**

Recall is used to measure the model's ability to correctly identify all data instances that truly belong to the positive class, thereby reflecting the model's effectiveness in retrieving relevant data.

$$\text{Recall} = \frac{TP}{TP+FN} \tag{3}$$

Keterangan :
TP = True Positive (TP) **(**Actual positive class that is correctly predicted as positive).
FN = False Negative (FN) **(**Actual positive class that is incorrectly predicted as negative).

- **F1-Score**

The F1-score is used to combine precision and recall values, thereby providing a more balanced and interpretable measure of model performance, particularly when the class distribution is imbalanced.

$$\textit{F1-Score} = 2 \times \frac{\textit{Precision x recall}}{\textit{Precision +recall}} \tag{4}$$

In this study, accuracy was calculated globally, whereas precision, recall, and F1-score were evaluated specifically for each comment category. Consequently, the values of TP, TN, FP, and FN vary for each sentiment class.

RESULT

The data collected through the web scraping technique resulted in 220 TikTok comments related to the issue of oil palm plantations and public figures in Indonesia. After data collection, the dataset underwent preprocessing, which included cleaning, stopword removal, tokenization, and stemming, followed by sentiment labeling using TextBlob. The sentiment labeling results using the TextBlob method indicate that out of a total of 220 data samples, 118 were labeled as neutral, 64 as positive, and 38 as negative. Subsequently, for the purpose of binary classification, the neutral class was assumed as the negative class, resulting in a final data distribution of 156 negative samples and 64 positive samples.

Table 7. Number of Sentiment Labels

Sentiment Label	Number of Comments
Positif	64
Negatif	156

Subsequently, the dataset was divided into 80% training data and 20% testing data, consisting of 176 and 44 comments, respectively. Next, a weighting technique using TF-IDF was applied to prepare the data for analysis using the K-Nearest Neighbors (KNN) algorithm in order to classify sentiments into positive, negative, and neutral categories. Based on the classification results, 31 data samples were classified as positive sentiment and 189 as negative sentiment.

Table 8. Sentiment Classification Distribution

Sentiment Classification	Number of Comments	Presentase (%)
Positive	31	14%
Negative	189	86%
Amount	220	100%

E-ISSN : 2807-9035
Volume 6, **Number** 1, May 2026
https://doi.org/10.47709/brilliance.v6i1.8073

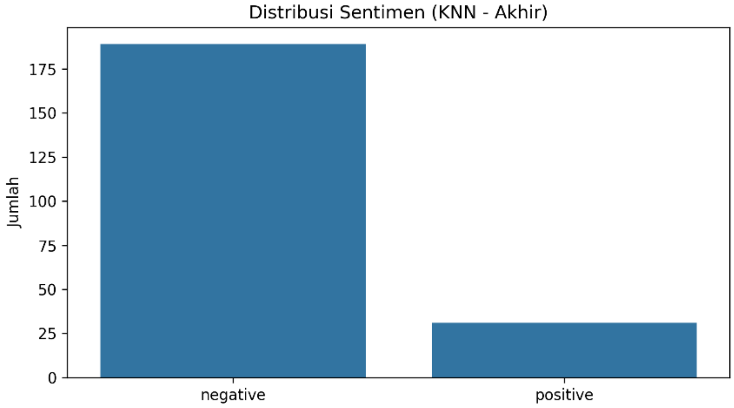

Figure 3. Bar Chart of Sentiment Distribution

Table 9. Classification

Tweet_class	Precision	Recall	F1-Score	support
Negative	0.7838	0.9355	0.8529	31
Positive	0.7143	0.3846	0.5000	13
Accuracy			0.7727	44
Macro avg	0.7490	0.6600	0.6765	44
Weighted avg	0.7633	0.7727	0.7487	44

Based on the testing results of the K-Nearest Neighbors (KNN) model, the performance evaluation metrics are presented in Table 9, including precision, recall, F1-score, and accuracy, which are used to measure the effectiveness of the model in classifying negative and positive sentiments.

For the negative class, the model achieved a precision of 0.7838, recall of 0.9355, and F1-score of 0.8529, with 31 test samples (support). The high recall value indicates that most of the truly negative data were correctly identified by the model. This demonstrates that the model has a strong capability in recognizing negative sentiment patterns.

Meanwhile, for the positive class, the model obtained a precision of 0.7143, recall of 0.3846, and F1-score of 0.5000, with 13 test samples. The relatively low recall value suggests that a considerable number of positive instances were misclassified as negative. This condition may be attributed to the class imbalance problem, where the number of negative samples significantly exceeds the number of positive ones, causing the model to be biased toward the majority class.

Overall, the model achieved an accuracy of 77.27%, indicating that most of the test data were classified correctly. The macro-average F1-score of 0.6765 reflects the average performance across both classes, while the weighted-average F1-score of 0.7487 shows that the model's performance is more influenced by the dominant class, namely the negative class.

Table 10. Confusion Matrix

	Pred_negative	Pred_positive
Actual_negative	29	2
Actual_positive	8	5

Based on the confusion matrix results, the model achieved a True Positive (TP) value of 5, True Negative (TN) of 29, False Positive (FP) of 2, and False Negative (FN) of 8. The evaluation results indicate that the model achieved an accuracy of 77.27%, a precision of 71.43%, a recall of 38.46%, and an F1-score of 50.1%. The relatively high accuracy suggests that the model is generally capable of performing classification effectively however, the relatively low recall indicates that a considerable number of positive samples are still misclassified as negative.

The following figure 4 presents the visualization of the confusion matrix, which illustrates the performance of the K-Nearest Neighbors (KNN) classification model in categorizing tweets into positive and negative sentiment classes. This confusion matrix displays the comparison between the actual labels and the predicted labels, thereby providing a clear representation of the classification accuracy and misclassification patterns produced by the model.

Brilliance
Research of Artificial Intelligence

E-ISSN : 2807-9035
Volume 6, **Number** 1, May 2026
https://doi.org/10.47709/brilliance.v6i1.8073

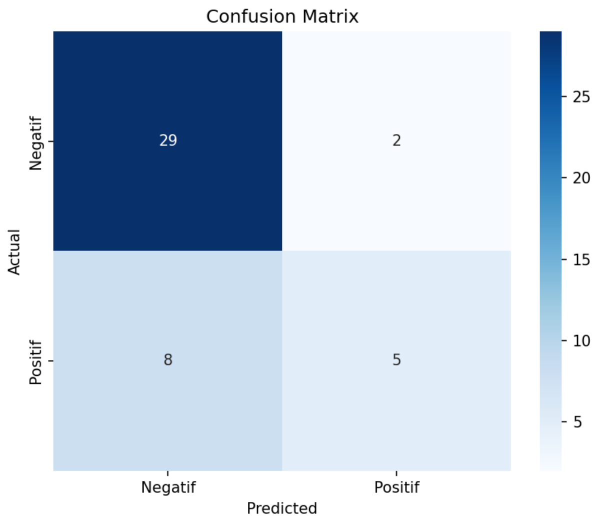

Figure 4. Confusion Matrix

DISCUSSION

Data were automatically collected from the TikTok platform using Python-based web scraping techniques. This process was conducted to obtain user comments related to the research topic, which were subsequently used as the dataset for sentiment analysis. A total of 220 comments were collected.

The preprocessing stage is a crucial step in sentiment analysis research, as the quality of the data significantly influences the performance of the classification model. Raw data obtained from social media platforms, particularly TikTok, generally contain various types of noise, such as symbols, emojis, URLs, abbreviations, and non-standard words. Therefore, data cleaning and normalization are required before further processing. The preprocessing stage aims to transform raw data into cleaner, more structured, and analysis-ready data for subsequent feature extraction and classification.

In this study, the preprocessing process not only included data cleaning and normalization but also incorporated a translation stage into English. This additional step was introduced because sentiment labeling using the TextBlob library yields more optimal results when applied to English text compared to Indonesian text. Consequently, all textual data that had undergone preprocessing were translated into English before the sentiment labeling process. Based on the sentiment labeling results using TextBlob, 118 samples were labeled as neutral, 64 as positive, and 38 as negative. Subsequently, for the purpose of binary classification, the neutral class was assumed to belong to the negative class. This assumption was based on the consideration that neutral sentiment tends not to reflect support or positive inclination toward the analyzed object. As a result, the final data distribution consisted of 156 negative samples and 64 positive samples.

After completing preprocessing, translation, and sentiment labeling, the next step was to divide the dataset into training data and testing data. This data splitting aims to train the classification model to learn sentiment patterns from the data and to evaluate the model's predictive capability on previously unseen data.

In this study, the dataset was divided using an 80% training and 20% testing ratio. This proportion was selected because it is commonly applied in machine learning research and is considered to provide a balance between sufficient training data for model learning and representative testing data for performance evaluation. With a total dataset of 220 samples, this resulted in 176 training samples and 44 testing samples.

The data splitting process was conducted using random sampling to avoid bias during training and testing. Additionally, the class proportions were considered to maintain a balanced distribution between positive and negative classes in both training and testing sets. This strategy ensures that the model can learn the characteristics of each class optimally and achieve more stable and accurate classification performance.

Following data splitting, feature weighting was performed using the TF-IDF method to transform textual data into numerical vectors. This method assigns weights to each term based on its frequency of occurrence and rarity across the entire document collection, thereby highlighting the most informative words for sentiment determination.

Based on the sentiment classification results using the K-Nearest Neighbors (KNN) algorithm, negative sentiment was found to dominate over positive sentiment. These results indicate that the majority of Indonesian netizens tend to oppose the plan for oil palm plantation development in Papua. This negative stance is influenced by concerns over environmental impacts, such as increased flood risks, ecosystem degradation, and deforestation, as previously observed in regions such as Aceh and Sumatra. Moreover, Papua, which is recognized for its exceptionally high biodiversity, is considered to require special protection from natural resource exploitation activities.

To measure the performance of the classification model, an evaluation was conducted using a confusion matrix

E-ISSN : 2807-9035
Volume 6, **Number** 1, May 2026
https://doi.org/10.47709/brilliance.v6i1.8073

and several performance metrics, namely accuracy, precision, recall, and F1-score. The evaluation results demonstrate that the model is capable of performing sentiment classification with a relatively high level of accuracy. However, the recall value obtained was comparatively lower than the other metrics.

The low recall value indicates that a number of data samples that should have been classified into a particular class were incorrectly predicted as belonging to another class. In the context of this study, the low recall suggests that the model has not yet achieved optimal performance in detecting all data instances that truly belong to a certain class, particularly the minority class. This limitation may be attributed to several factors, including class imbalance, limited training data, and the complexity of linguistic patterns used in social media comments.

The imbalanced data distribution, where negative samples significantly outnumber positive samples, causes the model to focus more heavily on learning the characteristics of the majority class. As a consequence, the model's ability to recognize the minority class is reduced, leading to lower recall values. In addition, variations in language style, the use of informal expressions, and sarcasm in social media comments further complicate the model's ability to accurately capture sentiment patterns.

Another factor contributing to the low recall is the limitation of TF-IDF-based feature representation, which only considers word frequency without capturing semantic context. As a result, the model struggles to interpret implicit meanings in sentences, particularly in comments that are ambiguous or context-dependent.

CONCLUSION

Based on the results of this study, it can be concluded that the sentiment analysis process of netizen comments regarding oil palm plantation development in Papua was successfully conducted through systematic stages, including data collection, preprocessing, translation, automatic labeling, feature weighting using TF-IDF, and classification using the K-Nearest Neighbors (KNN) algorithm. These stages form a comprehensive and effective methodological framework for processing social media text data.

The sentiment classification results reveal a dominance of negative sentiment over positive sentiment. This finding indicates that the majority of Indonesian netizens tend to oppose the plan for oil palm plantation development in Papua. This negative perception is primarily influenced by concerns over environmental impacts, such as increased flood risk, ecosystem degradation, and threats to biodiversity, as observed in other regions such as Aceh and Sumatra. These results highlight that environmental considerations are the main factors shaping public perception toward oil palm development policies.

The performance evaluation of the model using a confusion matrix and evaluation metrics demonstrates that the KNN algorithm is capable of performing sentiment classification with a relatively high level of accuracy. However, the recall value obtained is comparatively lower, indicating that the model still has limitations in detecting all data instances belonging to certain classes, particularly the minority class. This condition is influenced by class imbalance and the linguistic complexity of social media comments.

Overall, this study demonstrates that a machine learning–based sentiment analysis approach can be effectively utilized to capture public opinion on social and environmental issues. The findings are expected to serve as valuable input for policymakers in formulating more sustainable and environmentally conscious development strategies, particularly in the management of natural resources in Papua

REFERENCES

Alvionika, N., Faisal, S., Rahmat, R., & Masruriyah, A. F. N. (2024). Analisis Sentimen Pada Komentar Instagram Provider By.U Menggunakan Metode K-Nearest Neighbors (KNN). *Jurnal Algoritma*, *21*(2), 50–63. https://doi.org/10.33364/algoritma/v.21-2.1672

BAPPENAS, BRIN, & BPLH. (2024). *Status keanekaragaman hayati ekoregion sulawesi 2024*. 384. https://perpustakaan.bappenas.go.id/e-library/file_upload/koleksi/dokumenbappenas/konten/Dokumen 2025/Konten/Buku_Status Kehati Ekoregion Sumatra_.pdf

Hakim, B. (2021). ANALISA SENTIMEN DATA TEXT PREPROCESSING PADA DATA MINING DENGAN MENGGUNAKAN MACHINE LEARNING DATA TEXT PRE-PROCESSING SENTIMENT ANALYSIS IN DATA MINING USING MACHINE LEARNING School of Computer Science and Technology , Harbin Institute of Technology. *Journal of Business and Audit Information Systems*, *4*(2), 16–22.

Ayuk Pusvita, E., Stefanus, D., Ranggup, P., & Arfan, U. (2025). *MALCOM: Indonesian Journal of Machine Learning and Computer Science Machine Learning-Based Sentiment Analysis of Public Opinion on Palm Oil Plantation Expansion in Papua Implementasi Machine Learning untuk Analisis Sentimen Opini Publik Mengenai Pembukaan*. *5*(October), 1396–1405.

Indri Anggraini, & Prima Mulyasari Agustini. (2025). Tiktok sebagai Mesin Viralitas: Studi Tren Boost Juice di Kalangan Gen Z Jabodetabek. *Da'watuna: Journal of Communication and Islamic Broadcasting*, *5*(4), 387–401. https://doi.org/10.47467/dawatuna.v5i4.8179

Karima, A., Ulya, A., Sukiman, T. S. A., & Zufia, A. (2025). *Twitter Sentiment Analysis on the Iran-Israel Conflict*

E-ISSN : 2807-9035
Volume 6, **Number** 1, May 2026
https://doi.org/10.47709/brilliance.v6i1.8073

Using the Naïve Bayes Classification Algorithm. *6*(2), 930–943. https://doi.org/10.30596/jcositte.v6i2.26093

Li, C., Fu, J., Lai, J., Sun, L., Zhou, C., Li, W., Jian, B., Deng, S., Zhang, Y., Guo, Z., Liu, Y., Zhou, Y., Xie, S., Hou, M., Wang, R., Chen, Q., & Wu, Y. (2023). Construction of an Emotional Lexicon of Patients With Breast Cancer: Development and Sentiment Analysis. *Journal of Medical Internet Research*, *25*(1). https://doi.org/10.2196/44897

Murtiwiyati, M., Arini, R. F., Safitri, L., Widianti, S., & Setiadi, I. (2025). Implementasi Algoritma K-Nearest Neighbors untuk Klasifikasi Sentimen pada Ulasan Pakaian Anak di Platform TikTok Shop. *Jurnal Minfo Polgan*, *14*(2), 2921–2928. https://doi.org/10.33395/jmp.v14i2.15538

Nisa, A. D. M. K., Sa'diah, N. L., Fitriyani, L., Fitranti, A. N., & Setiawan, A. (2025). Penerapan Machine Learning Dengan K-Nearest Neighbors (Knn) Untuk Identifikasi Jenis Bunga Berdasarkan Karakteristik Fisik. *Jmsi*, *6*(2), 205–213.

Pinem, T. H., & Putra, Z. P. (2025). Evaluasi Kinerja Algoritma Klasifikasi Deep Learning dalam Prediksi Diabetes. *Jurnal Ilmiah FIFO*, *17*(1), 17. https://doi.org/10.22441/fifo.2025.v17i1.003

Purnamasari, D., Aji, A. B., Wulandari, D., Reza, F. A., Safrila, M., Yanda, N., & Hidayati, U. (2023). Pengantar Metode Analisis Sentimen. In *Pengantar Metode Analisis Sentimen*.

Rahmawati A. Damiti, Parid Pakaya, Mawardi Heru Prasetyo, Dewi Wahyuni K. Baderan, & Ramli Utina. (2025). Stabilitas Ekosistem Hutan Indonesia dalam Menghadapi Deforestasi dan Kerusakan Lingkungan: Tinjauan Literatur. *Botani : Publikasi Ilmu Tanaman Dan Agribisnis*, *2*(2), 176–188. https://doi.org/10.62951/botani.v2i2.343

Randy Suryono, R. (2025). Sentiment Analysis of the Influence of the Korean Wave in Indonesia Using the Naive Bayes Method and Support Vector Machine Analisis Sentimen Pengaruh Korean Wave Di Indonesia Dengan Metode Naive Bayes Dan Support Vector Machine. *Jurnal Inovtek Polbeng - Seri Informatika*, *10*(1), 308–319.

Sari, H., Leonarde Ginting, G., Zebua, T., & Mesran. (2021). Penerapan Algoritma Text Mining Dan TF-IDF untuk Pengelompokan Topik Skripsi pada Aplikasi Repository STMIK Budi Darma. *Terapan Informatika Nusantara*, *2*(7), 414–432. https://ejurnal.seminar-id.com/index.php/tin

Wong, J., Sanu, I., & Irsyad, H. (2025). Implementasi TF-IDF, Cosine Similarity, dan Logistic Regression Pada Rekomendasi Buku Berdasarkan Mood Pembaca Dengan Data Oversampling. *Device : Journal of Information System, Computer Science and Information Technology*, *6*(1), 142–154. https://doi.org/10.46576/device.v6i1.6499