Available *online* at: http://ajeeet.ft.unand.ac.id/

Andalas Journal of Electrical and Electronic Engineering Technology

ISSN 2777-0079

Sentiment Analysis of Traffic Congestion in Palembang Using Random Forest

Sukemi [1], Ahmad Fali Oklilas [2], Muhammad Azrell Samudra [3]

[1] *Department of Computer Systems, Universitas Sriwijaya, Srijaya Negara Street, Palembang, 30139, Indonesia*
[2] *Doctoral Student in Engineering Science, Faculty of Engineering, Universitas Sriwijaya, Srijaya Negara Street, Palembang, 30139, Indonesia.*
[3] *Department of Computer Systems, Universitas Sriwijaya, Srijaya Negara Street, Palembang, 30139, Indonesia*

ARTICLE INFORMATION

Received: March 20, 2026
Revised:April 02, 2026
Accepted: May 07, 2026
Available online: May 31, 2026

KEYWORDS

Sentiment Analysis, Random Forest, TF-IDF, Facebook, Traffic Congestion

CORRESPONDENCE

Phone: +62 87775122044
E-mail: sukemi@ilkom.unsri.ac.id

A B S T R A C T

Traffic congestion is a persistent problem that significantly affects the daily activities of citizens in Palembang City. The public can voice their thoughts and worries about traffic conditions on social media, especially Facebook. This study uses the Random Forest algorithm to examine public opinion regarding traffic congestion in Palembang. The 2,021 Facebook comments in the dataset were gathered by web scraping and subjected to a number of preprocessing steps, such as cleaning, case folding, stemming, tokenization, normalization, and stopword removal. The TF-IDF algorithm was used for term weighting. The Random Forest model was trained and tested to classify sentiments into three categories: positive, neutral, and negative. The model attained good accuracy across training, testing, and validation datasets, according to the evaluation results. This research provides insights into public perceptions of traffic congestion and can serve as a reference for policymakers in developing data-driven strategies to address traffic issues in Palembang City.

INTRODUCTION

Many major Indonesian cities, including Palembang, struggle with traffic congestion. Congestion has grown to be a complicated issue that interferes with the community's everyday activities due to the increase in the number of vehicles and the constraints of the road infrastructure[1]. The impact of traffic congestion is not limited to longer travel times but also contributes to air pollution, increased vehicle operating costs, and reduced public productivity[2].

In the current digital age, people frequently utilize social media sites like Facebook to voice their thoughts and grievances about a variety of problems, including traffic jams[3]. Social media posts and comments can offer insightful information about how the general public feels about this subject.

Sentiment analysis is one of the methods that can be used to extract such information, with the aim of gaining a deeper understanding of public views[4]. Processing large amounts of social media data requires effective and efficient algorithms.

One machine learning model that has shown improved performance in classification and prediction based on complex data is the Random Forest algorithm. This algorithm works by constructing multiple decision trees and producing a final classification based on the voting results from each tree. Therefore, it is capable of delivering more accurate results in sentiment analysis[5]. In the context of urban management, understanding public sentiment is increasingly recognized as an important component in supporting evidence-based decision making. Public opinions expressed through social media often reflect real conditions experienced by road users, such as delays, road bottlenecks, parking violations, and traffic management issues. Unlike traditional survey methods, social media analysis enables the collection of large-scale data in a relatively short time and at a lower cost, making it a valuable alternative source of information for urban studies.

Palembang City, as one of the major urban centers in Sumatera Selatan, faces persistent traffic congestions problems, particularly in areas with high economic activity, educational institutions, and commercial centers. Peak-hour congestion frequently occurs due to a combination of factors such as limited road capacity, increasing private vehicle usage, roadside parking, and inadequate traffic regulation enforcement. Social media platforms such as Facebook have become an important medium for people to express their opinions regarding public issues, including traffic congestion. In Indonesia, Facebook remains one of the most widely used social media platforms across different demographic group. Therefore, comments posted on Facebook

https://doi.org/10.25077/ajeeet.v6i1.214

can provide valuable insights into public perceptions and experiences related to traffic conditions.

Sentiment analysis, also known as opinion mining, plays a crucial role in extracting subjective information from textual data. By classifying opinions into categories such as positive, neutral, and negative, sentiment analysis enables researchers to quantitatively measure public attitudes toward specific issues[6]. In the case of traffic congestion, sentiment classification can help identify whether the public response tends to be dominated by dissatisfaction, acceptance, or positive perceptions related to traffic conditions or management efforts. This information is valuable for evaluating existing transportation policies and identifying areas that require improvements[7].

Previous studies have applied sentiment analysis techniques to analyze public opinions related to urban issues using social media data. Several machine learning algorithms such as Naïve Bayes, Support Vector Machine, and decision tree-based models have been widely used for sentiment classification tasks. However, these methods may face limitations in handling complex patterns within large-scale textual data. Therefore, this study applies the Random Forest algorithm to improve classification robustness and performance.

This research is expected to provide a clearer understanding of public perceptions of traffic congestion and contribute to policymakers in formulating more effective solutions to address the issue. Through this study, it is hoped that a better understanding of the data available on social media can assist the government and relevant stakeholders in formulating more targeted, data-driven policies to mitigate traffic congestion in Palembang City.

METHOD

Data Gathering Techniques

The data collection process was carried out using a crawling technique through the Apify Console to retrieve comments from Facebook social media. In this process, the Facebook Comments Scraper was used. The URL of the Facebook post was entered into the actor for comment scraping, after which the actor was executed (start actor) to generate raw data that could then be exported in CSV format[8]. Figure 1 shows the workflow for data collection.

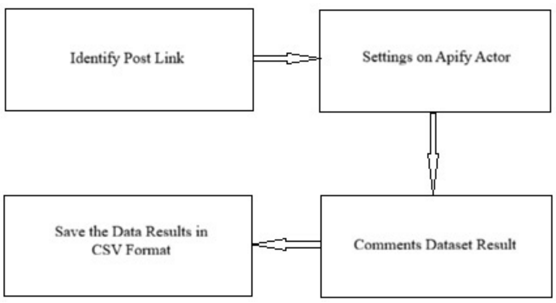

Figure 1. Data Collection Process

The raw data obtained through the web scraping process was subsequently used as the primary material for sentiment analysis. The selection of Facebook as the data source is considered appropriate because the platform has a very large number of active users in Indonesia and is frequently used by the public to express opinions on various public issues, including traffic congestion[9]. The scraping results were then stored in CSV (Comma Separated Values) format to facilitate the next processing stage using the Python programming language. Each entry in the file contains the comment text, upload time, and anonymized account identity.

Preprocessing Data Text

Preprocessing was done to get the text data ready for machine learning algorithms to process it efficiently. Raw data obtained from web scraping generally still contains various unnecessary elements, such as punctuation marks, emojis, links, and words that do not provide informative value. In addition, the raw ext often includes inconsistent writing styles, spelling variations, and noise caused by user-generated content, which may negatively affect the performance of the classification model if not handled properly. Although social media comments may contain subjective opinions, they still reflect real experiences shared by users. To maintain data validity, several filtering procedures were applied, including the removal of duplicate comments, spam content, and irrelevant text. Only comments related to traffic congestion in Palembang City were retained for further analysis. Therefore, text-cleaning procedures are required to produce more structured data suitable for use in sentiment analysis[10]. Figure 2 depicts the preprocessing workflow.

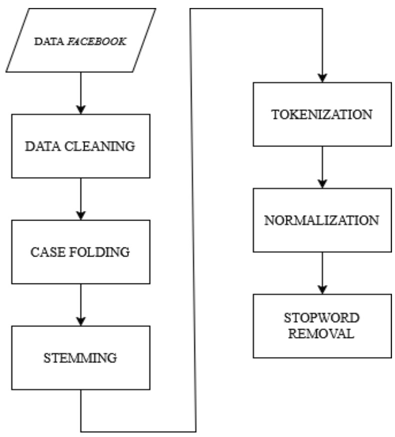

Figure 2. Preprocessing Flow for Text Data

Data cleaning, case conversion, stemming, tokenization, normalization, and stopword removal were the pre-processing steps used in the study.

Data Cleaning

Data Cleaning is the initial preprocessing step aimed at removing irrelevant and noisy elements from raw text data. For sentiment analysis, data cleaning entails removing punctuation, numerals, emoticons, URLs, special characters, and excessive whitespace. This stage is crucial for lowering dataset noise and guaranteeing

 https://doi.org/10.25077/ajeeet.v6i1.214

that only pertinent textual material is kept for additional processing.

Case Folding

Next step is case folding, that applied to standardize all text into lowercase form, such as *"Macet"*,*"MACET",* and *"macet"*, which represent the same word but may be treated as different tokens by machine learning algorithms.

Stemming

Stemming is the process of returning words to their root or basic form according to the rules of Indonesian morphology. This step aims to unify different word variations that share the same meaning, such as converting *"berjalan","berjalannya",*and *"perjalanan"* into their root forms. By applying stemming, redundant word forms are minimized.

Tokenization

Tokenization is the next step after stemming, that breaking text into smaller units called tokens, typically individual words. Each comment is split into sequence of word to enable further analysis and feature extraction.

Normalization

The next step is normalization that performed to standardize non-standard, informal, or slang words into their formal equivalents without changing their original meanings. For examples, word such as *"gak"* are converted into *"tidak"*, and *"macetnyo"* into *"macet"*.

Stopword Removal

And the last step is stopword removal, the process of eliminating commonly used words that frequently appear in text but do not contribute significantly to sentiment determination. The Indonesian terms "dan," "yang," "di," and "ke" are examples of stopwords. Eliminating these phrases improves the classification process' efficacy by lowering dimensionality and enabling the model to concentrate on more significant terms that convey sentiment information[11][12][13].

Filtering

Next, the filtering stage is carried out to ensure that only relevant and valid data are used for analysis. At this stage, duplicate texts, empty comments, and data containing non-text elements such as links, numbers, or special characters are removed to improve dataset quality before the weighting and classification processes. Filtering also includes checking text length, as comments that are too short or too long may introduce bias into the sentiment analysis results[14]. This highlights the significance of data screening because extraneous text may lower the machine learning model's accuracy. Besides that, contextual verification is performed to ensure that each comment is truly related to the issue of traffic congestion in Palembang City.

Splitting Data

In the data splitting stage, the dataset that has undergone preprocessing and filtering is divided into three subsets, namely the training set, validation set, and testing set. The purpose of these three subsets is to make sure the model can both learn patterns from the training data and be gradually assessed before being put to the test with fresh, unprecedented data[15].

The data was divided into training, testing, and validation sets using the stratified hold-out method, which maintains the percentage of each class to ensure balanced feature distribution. To apply this method, 70% of the dataset was utilized for training, 20% for testing, and 10% for validation.

TF - IDF *(Term Frequency – Inverse Document Frequency)*

In the weighting stage, the TF - IDF (Term Frequency–Inverse Document Frequency) method is used to convert the filtered text into a numerical form that can be processed by classification algorithms[16]. TF - IDF assigns a weight to each word based on its frequency of occurrence in a document relative to the entire corpus. Words that appear frequently in one document but are rarely found in other documents will receive a higher TF - IDF weight[17][18].

$$TF-IDF(t,d) = TF(t,d) \times log\left(\frac{N}{df(t)}\right) \tag{1}$$

In this formula, *tf* represents the frequency of occurrence of word *t* in a document, *df* indicates the number of documents containing word *t*, while *N* refers to the total number of documents in the entire dataset. The TF value reveals how frequently a term appears within a single document, but the IDF number shows how uncommon a word is across all texts[19]. By combining these two values, TF - IDF is able to emphasize the most relevant or representative words in a comment. This method is effective because it reduces the dominance of common words that do not carry specific meaning.

The TF - IDF weighting scheme ensures that words which frequently appear in a particular document but rarely occur in the overall corpus are assigned higher weights. Consequently, these words are considered more informative and discriminative for classification tasks, including sentiment analysis. On the other hand, words that appear frequently across many documents, such as general terms or stopwords, receive lower weights, thereby minimizing their influence on the classification process[20].

In the context of social media text analysis, the application of TF-IDF is particularly important due to the informal and diverse nature of user-generated content. Social media comments often contain repetitive expressions, filler words, and common terms that do not contribute significantly to sentiment polarity. By applying TF - IDF, the textual data can be transformed into a numerical representation that highlights meaningful sentiment-bearing words while suppressing less informative ones[18].

Random Forest Algorithm Classification

The Random Forest method is then used for sentiment classification. This algorithm is a supervised learning technique that builds several decision trees and then aggregates the predicted outcomes from each tree. The Random Forest algorithm was selected because of its ability to handle high-dimensional data and reduce the risk of overfitting. By combining multiple decision tress, Random Forest can produce more stable and accurate predictions compared to single-tree classifiers. This characteristic makes it suitable for sentiment analysis tasks involving textual data represented using TF-IDF features. Random Forest utilizes the attributes and feature values obtained from the TF-IDF weighting to determine sentiment categories. Compared to employing a single decision tree, this ensemble approach allows the model to generate predictions that are more reliable and accurate[21][22]. Figure 3 provides an illustration of this process.

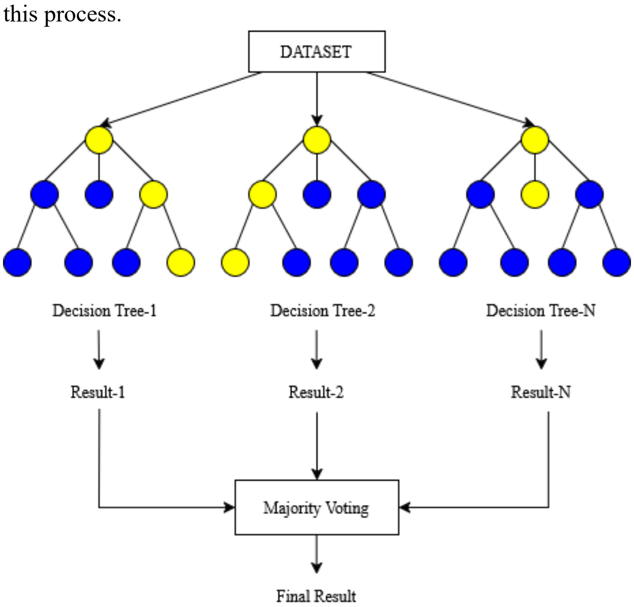

Figure 3: Classification Model Using Random Forests

In the Random Forest approach, each decision tree consists of nodes that represent attributes, branches that display the results of attribute testing, and leaf nodes with sentiment class labels. Unlike a single decision tree, Random Forest constructs many trees simultaneously and then combines the prediction results from all trees to produce a final decision that is more accurate and stable[23].

The primary benefit of Random Forest is its capacity to enhance classification performance using an ensemble method, which makes it more resilient to overfitting and able to handle big datasets and intricate features like TF-IDF numerical representations. In addition, this method still provides interpretability through feature importance information, which indicates which words or features have the most influence in the classification process.

The Random Forest model is employed in this work to categorize public comments into three sentiment groups: neutral, negative, and positive. The training and testing data are used to train and evaluate the model's ability to generalize. Sentiment labels reflecting public opinions on Palembang City's traffic congestion problems make up the stage's output.

RESULTS AND DISCUSSION

Data Collection Results

The data in this study were obtained through a web scraping process of public comments on Facebook discussing traffic congestion issues in Palembang City. The data were collected from several posts on local news pages as well as community forums that actively discuss traffic conditions. Because it allows for the automatic retrieval of vast amounts of data from online sites, the web scraping approach was chosen in order to save the longer time needed for human data collecting. Table 1 provides an illustration of this procedure.

Table 1. Facebook Unprocessed Data

Date	Time	Text
11/4/2024	8:30:24	*Ya parahlah kitu kalo tak disiplin*
11/5/2024	0:12:51	*Gak ada tukang parkir kah?? Lumayan itu bisa jadi kerjaan*
11/4/2024	5:26:30	*Lampu merah itu jugo kadang aneh merah galo kadang ijo galo jdi ktk fungsi*
9/10/2024	1:58:28	*Saat hujn jadi kebanyakan bawa mobil*
12/11/2024	3:11:43	*Bimbel BTA, SDIT Alfurqon, dan ojol mobil motor parkir dibahu jalan depan PTC itulah biang macetnyo*

The raw data obtained were stored in Comma Separated Values (CSV) format. At this stage, data collection was carried out by entering the keywords *"palembang"* and *"macet."* The resulting dataset consists of 5.842 rows of data.

Text Data Preprocessing Results

The preprocessing and filtering stages were carried out to prepare the collected comment data so that it is suitable for use in sentiment analysis. This process includes a series of text-cleaning steps, such as removing punctuation marks, numbers, URLs, excessive spaces, and non-alphabetic characters, with the aim of eliminating elements that do not hold contextual meaning. In this study, the preprocessing stages include case folding, stemming, tokenization, normalization, and stopword removal. These results can be seen in Table 2.

Table 2. Example of Text Data Preprocessing

No.	Text	Tokenized	Normalization
1	*Ya parahlah kitu kalo tak disiplin*	*['ya', 'parah', 'kitu', 'kalo', 'tak', 'disiplin']*	*ya parah kitu kalo disiplin*
2	*Gak ada tukang parkir kah?? Lumayan itu bisa jadi kerjaan*	*['gak', 'ada', 'tukang', 'parkir', 'kah', 'lumayan', 'itu', 'bisa', 'jadi', 'kerjan']*	*enggak tukang parkir kah lumayan kerjan*
3	*Lampu merah itu jugo kadang aneh merah galo*	*['lampu', 'merah', 'itu', 'jugo', 'kadang', 'aneh', 'merah',*	*lampu merah jugo kadang aneh merah galo*

	kadang ijo galo jdi ktk fungsi	*'galo', 'kadang', 'ijo', 'galo', 'jdi', 'ktk', 'fungsi']*	*kadang ijo galo jadi ktk fungsi*
4	*Saat hujn jadi kebanyakan bawa mobil*	*['sat', 'hujn', 'jadi', 'banyak', 'bawa', 'mobil']*	*sat hujn bawa mobil*

Table 2 presents the results of the tokenization and normalization processes applied to public comments from Palembang regarding traffic congestion. The tokenization process produces a sequence of words that are ready to be used in the weighting stage, while normalization aims to standardize non-standard words into consistent forms without altering their meanings. During the normalization stage, repeated words and informal expressions were standardized to their base forms. For instance, repeated expressions such as "*macet semacet macetnya*" were simplified to the root word "*macet*". In addition, negation expressions such as "*tidak macet*" or "*belum macet*" were preserved during preprocessing to ensure that contextual meaning of the sentiment was not altered.

Filtering Results

Following preprocessing, a filtering procedure is carried out to make sure the examined data actually corresponds with the study's objective, which is the public's views of traffic congestion in Palembang City. At this stage, irrelevant comments such as promotions, spam, or conversations unrelated to traffic issues are removed from the dataset. This process is illustrated in Table 3.

Table 3. Text Data Filtering

Normalization	Class
lampu merah jugo kadang aneh merah galo kadang ijo galo jadi ktk fungsi	neutral
bimbel bta sdit alfurqon ojol mobil motor parkir bahu jalan ptc biang macetnyo	negative
nya jam malam lengang tidak macet	positive
kalo dari arah polda ptc galk macet kalo sebaliknyo les min parkir sembarang bikin macet men dari situ lengang ruponyo bikin macet be yang les bemobil galo parkr bahu jalan po indak cocok nian	negative
parah macet nya skip ujung nya plebaran jalan pasar skip ujung flyover	negative

From the total of 5,843 comments collected in the initial stage, 3,822 comments were eliminated because they did not match the context of the study, resulting in 2,021 clean data entries ready for analysis. Following the validation of the data, a manual sentiment labeling procedure was used to categorize the comments into three groups: neutral, negative, and positive. The general substance and context of the words as well as the expression of opinions in the text such as the use of terms that denote happy emotions (like "lancar") or negative emotions (like "macet") were taken into consideration when determining these classifications.

Data Splitting Results

The data splitting process was carried out after the filtering and sentiment labeling stages were completed. The purpose of this division is to separate the dataset into several subsets so that the machine learning model can be trained, validated, and evaluated objectively. The study's dataset was divided into three parts: 70% for training, 10% for validation, and 20% for testing.

1.414 of the 2,021 comments that were found to be valid were used as training data, 404 as testing data, and 203 as validation data. This division was carried out using the stratified hold-out method, a data separation technique that maintains the proportion of each class to ensure that the representation of each cluster remains balanced across all data subsets.

TF - IDF *Results*

An example of text data utilized in the computation of TF-IDF weights is shown below. The weighting calculation for each word is carried out by first determining the TF value, then calculating the IDF, and finally multiplying the two values. The weighting procedure can be applied to the sample text displayed in Table 4 using this formula.

Table 4. Data for Training

No.	Text
1	*bknyo diterapke gagal bknyo ngatasi macet macetnyo*
2	*macet mobil pribadi nyemput sekolah al furqon samo mengantri bm spbu*
3	*iyo pas jam malam teraso lancar*
4	*kalo dari arah polda ptc galk macet kalo sebaliknyo les min parkir sembarang bikin macet men dari situ lengang ruponyo bikin macet be yang les bemobil galo parkr bahu jalan po indak cocok nian*
5	*enggak macet yang bangun tol jalan nasional kota jalan nasional propinsi kabupaten enggak lebar kota sudah jalan tol yang lebar bagus*

In the first sentence in Table 3, "*bknyo diterapke gagal bknyo ngatasi macet macetnyo,*" there are 7 words. TF determines how frequently each word occurs in the text. For instance, if the word "macet" appears twice, the TF computation is as follows.

$$\text{TF} = \frac{numbers\ of\ t\ in\ d}{total\ words\ in\ d} = \frac{2}{7} = 0{,}285$$

Thus, the TF value for the word "macet" in document 1 is 0.285. Meanwhile, the IDF calculation for the word "macet," which appears in 10 documents in Table 2, is as follows:

$$idf_t = ln\frac{D+1}{Df_t+1} + 1 = ln\frac{10+1}{7+1} + 1 = ln\left(\frac{11}{8}\right) + 1 = 1{,}318$$

Following the acquisition of the TF and IDF values, the TF-IDF weight is computed by multiplying the TF and IDF outcomes as follows.

$$Wt, d = 0{,}285 \times 1{,}318 = 0{,}375$$

Thus, the TF - IDF value for the word "macet" is 0.375.

The use of TF-IDF in this study is supported by previous research emphasizing the importance of feature extraction in sentiment analysis. Comparison of several feature extraction techniques [24], including Bag-of-Words, Word2Vec, N-gram, TF-IDF, Hashing Vectorizer, and GloVe, and found that feature extraction plays a crucial role in improving machine learning-based sentiment classification. In addition, reference [25] has applied TF-IDF in social media sentiment analysis and showed that TF-IDF can effectively represent textual data for classification tasks. These findings support the present study, which used TF-IDF to transform Facebook comments into numerical features for processing by the Random Forest algorithm. Therefore, TF-IDF is appropriate for handling social media comments because it emphasizes important terms while reducing the influence of common or less informative words.

Random Forest Algorithm Results

After completing the preprocessing and data splitting stages, the next step is to train the model using the Random Forest algorithm. In this phase, several parameters are configured to achieve optimal model performance. The number of decision trees to be built is specified by the *n*_estimators argument, and the maximum depth of each tree is controlled by the max_depth parameter to avoid overfitting. The criterion parameter is left at its default setting, allowing the algorithm to automatically determine the node-splitting measure based on the training data. Furthermore, Bootstrap is used to guarantee that a randomly selected portion of the data is used to train each tree. With this configuration, RF can provide a reliable model that is resistant to overfitting and has better generalization skills.

The summary containing information on the accuracy achieved at each stage, the amount of data used, and its percentage relative to the entire dataset can be seen in Table 5**.**

Table 5. Summary of Training, Testing, and Validation Results

No	Data Type	Accuracy (%)	Total Data	Percentage (%)
1	Training	87,19	1.414	70
2	Testing	81,18	404	20
3	Validation	81,28	203	10

Table 5 shows that the Random Forest model's accuracy on the training set, which included 1.414 data points, or around 70% of the whole dataset, was 87.19%. In the testing phase, the model obtained an accuracy of 81.18% using 404 data points (20%). Meanwhile, the validation process resulted in an accuracy of 81.28% with a total of 203 data entries (10%).

The Random Forest model can consistently generalize to new data, as evidenced by the comparatively minimal accuracy discrepancies between the three stages. The model fits the data well, meaning it is neither overfitting nor underfitting, as evidenced by the consistent accuracy over the training, testing, and validation phases with just minor fluctuations.

The accuracy obtained in this study is comparable to previous sentiment analysis studies that applied Random Forest and other machine learning algorithms to social media data. In paper [26], Naïve Bayes, Decision Tree, and Random Forest have been compared in Twitter-based sentiment analysis and showed that Random Forest can be used effectively for classifying public opinion in Indonesian social media contexts. Similarly, in paper [24] it has been shown that Random Forest combined with appropriate feature extraction techniques can produce strong sentiment classification performance. In the present study, the testing accuracy of 81.18% and validation accuracy of 81.28% indicate that the Random Forest model maintained stable performance on unseen data. This result supports previous evidence that Random Forest is a robust ensemble-based classifier for sentiment analysis because it combines multiple decision trees to improve prediction stability and reduce overfitting.

Table 6. Evaluation Results of the Random Forest Model on Training Data.

Actual/Predicted	Predicted Negative	Predicted Neutral	Predicted Positive
Actual Negative	512	34	1
Actual Neutral	126	657	2
Actual Positive	13	5	64

Table 7. Classification Report of Training Data.

Classification Report	Precision	Recall	F-1 Score
Negative	0.79	0.94	0.85
Neutral	0.94	0.84	0.89
Positive	0.96	0.78	0.86

Table 6 shows that the model performs effectively in all three data subsets based on the training data evaluation findings. On the training dataset, the model achieved an accuracy of 87.19%, with a precision of 0.79. The recall values for the "negative" sentiment reached 0.95, for the "neutral" sentiment 0.94, and the corresponding F1-scores were also high, demonstrating that the model can successfully identify patterns in the training set.

However, the recall for the positive class was relatively low at 0.78, suggesting the presence of data imbalance. The following section presents a detailed explanation of the results.

Diagonal:

- TP_negative = 512
- TP_neutral = 657
- TP_positive = 64

True Total:

$TP_{total} = 512 + 657 + 64 = 1233$

Sample Total: $N = 1414$

Therefore:

$$Accuracy = \frac{\sum i TP_i}{N} \quad (2)$$

$$Accuracy = \frac{1233}{1414} = 0.8719 = 87.19\%$$

For I class:

- $Precision_i = \frac{TP_i}{(TP_i + FP_i)}$
- $Recall_i = \frac{TP_i}{(TP_i + FN_i)}$
- $F1_i = 2 \cdot \frac{Precision_i \cdot Recall_i}{(Precision_i + Recall_i)}$

Numeric Results:

Negative:

- $Precision = \frac{512}{(512+139)} = 0{,}7864 = 0.79$

 https://doi.org/10.25077/ajeeet.v6i1.214

- $Recall = \frac{512}{(512+35)} = 0{,}9360 = 0.94$
- $F1 = 0.85$

Neutral:

- $Precision = \frac{657}{(657+35)} = 0{,}9494 = 0.94$
- $Recall = \frac{657}{(657+128)} = 0{,}8433 = 0.84$
- $F1 = 0.89$

Positive:

- $Precision = \frac{64}{(64+3)} = 0{,}9552 = 0.96$
- $Recall = \frac{64}{(64+18)} = 0{,}7804 = 0.78$
- $F1 = 0.86$

Table 8. Evaluation Results of the Random Forest Model on Testing Data.

Actual/Predicted	Predicted Negative	Predicted Neutral	Predicted Positive
Actual Negative	134	11	0
Actual Neutral	50	185	0
Actual Positive	10	5	19

Table 9. Classification Report of Testing Data.

Classification Report	Precision*	Recall**	F-1 Score***
Negative*	0.69	0.92	0.79
Neutral**	0.92	0.79	0.85
Positive***	1.00	0.38	0.55

Based on Table 5, the accuracy of the test data decreased slightly to 81.18%, which is normal because the model is faced with new data that has not been seen before. This fact shows that the model does not only memorize the training data patterns, but tries to generalize to different data.

Although the accuracy value is still quite good, evaluation of other metrics also needs to be considered, especially in the positive class in Table 9. The low positive class recall value (0.38) indicates that the model is still not able to identify all positive data optimally. Data imbalance (imbalanced data), which results from the quantity of data in the negative class being significantly more prevalent than the positive class, is the source of this problem.
This imbalance causes the model to tend to focus more on learning patterns from the majority class, so that most of the minority class data is not detected properly.

This finding is consistent with previous studies showing that class imbalance can reduce model performance in minority sentiment categories. Referring to paper [27] explains that imbalanced datasets remain a major challenge in sentiment classification because models tend to learn more strongly from majority classes and may fail to recognize minority classes accurately. Similarly, reference [28] has reported that unbalanced sentiment distributions can affect classification performance and requires specific strategies to improve minority class detection. In the present study, the positive sentiment class showed a lower recall value than the negative and neutral classes, indicating that several positive comments were not correctly identified by the model. Therefore, although the overall testing accuracy was relatively good, the low recall value for positive sentiment suggests that accuracy alone is not sufficient to evaluate model performance when the dataset is imbalanced.

In Table 5, the accuracy value on the validation data was recorded at 81.28%, which shows that the model was able to maintain relatively good performance when tested on data that was different from the training data. This accuracy value also does not experience a significant difference compared to the accuracy on the test data, so it can be concluded that the model does not experience overfitting and has a fairly stable generalization capability.

Table 10. Evaluation Results of the Random Forest Model on Validation Data.

Actual/Predicted	Predicted Negative	Predicted Neutral	Predicted Positive
Actual Negative	66	5	0
Actual Neutral	26	92	1
Actual Positive	4	2	7

Table 11. Classification Report of Validation Data.

Classification Report	Precision*	Recall**	F-1 Score***
Negative*	0.69	0.93	0.79
Neutral**	0.93	0.77	0.84
Positive***	0.88	0.54	0.67

If we look further at the evaluation per class based on Table 11, the recall value for the positive class in the validation data has increased to 0.54 compared to the results in the test data. This improvement indicates that the model is starting to show better ability in recognizing some of the data that falls into the positive class. However, the recall value is still moderate and not optimal, which indicates that the problem of imbalanced data still affects the model's performance, especially in classifying minority classes.

The validation result also supports previous evidence that Random Forest can maintain stable overall performance, although minority-class classification may still require improvement. The validation accuracy of 81.28% was very close to the testing accuracy of 81.18%, indicating acceptable generalization ability. However, the moderate recall value for the positive class suggests that the model still faced difficulty in identifying less frequent sentiment categories. Dataset characteristics, language variations, and evalution metrics are important challenges in social media-based sentiment analysis [29]. In line with this, Referring to previous papers [27] and [28] it has been suggested that imbalance-handling strategies are important when the distribution of sentiment classes is unequal. Future studies may therefore apply oversampling, undersampling, class weighting, or transfer learning approaches to improve the model's sensitivity toward positive sentiment data.

Analysis of Prediction Accuracy of the Random Forest Model on the Testing Data

To ensure that the Random Forest classification results on the testing data were accurate, each comment was manually reviewed, taking into account the overall meaning of the sentence as well as the context of the previous post. This step is necessary because some comments literally contain words like *"lancar"* but

https://doi.org/10.25077/ajeeet.v6i1.214

in reality, they refer to efforts to address traffic congestion in a particular location. Therefore, Table 12 includes a column for Previous Post to provide additional context during the verification process.

Table 12. Sentiment Classification Accuracy Analysis

Text	Real Label	Prediction	Truth Value
dari kemarin malam nian macet nya	negative	negative	1
bikin indak nyaman penguna jalan	negative	neutral	0
macet veteran	negative	negative	1
macet parah ya bun	negative	negative	1
muternyo tutup mangkonyo numpuk puter yang macet	negative	negative	1
sekip macet terus kak	negative	negative	1
jalan sempit dan gelombang mobil yang lewat padat truk tronton trailer bis besak dulu pokoknye dk kenal waktu macet	negative	negative	1
jalan rusak galo macet aduh palembang	negative	negative	1
macet arah tanjung api	negative	negative	1

Based on Table 12, the prediction results of the Random Forest model show a total of 404 test data instances divided into three sentiment categories: 145 negative, 235 neutral, and 24 positive instances. The model's overall accuracy of 81.18% shows that the majority of predictions match the labels. However, for the positive category, the model appears less capable of correctly classifying the data due to the presence of data imbalance.

Overall, the findings of this study are in line with previous research on machine learning-based sentiment analysis. Prior studies have shown that Random Forest is effective for sentiment classification because it uses an ensemble structure that combines multiple decision trees and improves classification stability[26], [29], [30]. The use of TF-IDF also strengthens the classification process by converting textual data into weighted numerical features that represent the importance of terms in the dataset[24], [25]. However, unlike studies with more balanced datasets, the present study faced challenges in classifying positive sentiment because the number of positive comments was smaller than the negative and neutral comments. This condition explains why the model achieved good overall accuracy but lower recall for the positive class. Thus, this study contributes to the existing literature by demonstrating that the TF-IDF and Random Forest approach can be applied to Facebook-based traffic congestion sentiment analysis in Palembang City, while also highlighting the need to address class imbalance in future studies.

CONCLUSIONS

This study has demonstrated that the Random Forest algorithm can be effectively applied for sentiment analysis of traffic congestion in Palembang City based on Facebook comments. The dataset consisted of 2,021 comments that were processed through comprehensive preprocessing stages and represented using the TF-IDF weighting method. The experimental results indicate that the proposed model achieved satisfactory performance, with accuracy values of 87.19% on the training data and 81.18% on the testing data, showing that Random Forest is capable of capturing sentiment patterns in social media text.

However, the evaluation also revealed that the model's performance on the positive sentiment class was relatively lower compared to the negative and neutral classes. This limitation is primarily caused by data imbalance, where the number of positive samples was significantly smaller than the other sentiment categories. Despite this limitation, the model was able to classify the majority of data correctly, indicating its robustness for sentiment classification tasks.

All things considered, this study offers insightful information on how the general public views traffic congestion in Palembang City and emphasizes the potential of sentiment analysis based on machine learning as a tool to assist policymakers in making decisions. Future research is recommended to address data imbalance issues by applying resampling techniques or exploring alternative classification methods to further improve model performance.

Another important difference is the classification algorithm method. Previous research has often used the KNN algorithm, which classifies data based on the proximity between data points. Although this method is simple and easy to implement, KNN has drawbacks in handling large datasets and is susceptible to noise. This study utilized the decision tree algorithm, which is capable of producing a classification model that is easier to interpret through a decision tree structure and is able to understand hierarchical data patterns based on the most influential attributes. The use of decision trees has proven to provide a more transparent and effective analytical approach, especially in the context of long text comments.

One limitation of this study is that the analysis relies solely on data collected from social media comments. Although social media provides large volumes of public opinion data, it may not fully represent the perspectives of all members of society. Future research may combine social media sentiment analysis with field surveys or questionnaires to obtain a more comprehensive understanding of public perceptions regarding traffic congestion. In addition, public opinions expressed on social media are dynamic and may change over time depending on various factors such as traffic conditions, user activity, and social trends. Continuous monitoring and analysis of social media data are therefore necessary to capture evolving public statement.

ACKNOWLEDGMENT

Acknowledgements are extended to the research group of Integration of Pervasive Computing, IoT, and Embedded Systems at the Computer Systems Department, Faculty of Computer Science, Universitas Sriwijaya, for the support provided, which made this research possible to be carried out successfully.

REFERENCES

[1] G. P. Rocha Filho *et al.*, "Enhancing intelligence in traffic management systems to aid in vehicle traffic congestion problems in smart cities," *Ad Hoc Networks*, vol. 107, 2020, doi: 10.1016/j.adhoc.2020.102265.

[2] M. Pi, H. Yeon, H. Son, and Y. Jang, "Visual Cause Aalytics for Traffic Congestion," *IEEE Trans. Vis. Comput. Graph.*, vol. 27, no. 3, pp. 2186–2201, 2021, doi: 10.1109/TVCG.2019.2940580.

[3] Ö. Çoban, S. A. Özel, and A. Inan, "Deep Learning-based Sentiment Analysis of Facebook Data: The Case of Turkish Users," *Comput. J.*, vol. 64, no. 3, pp. 473–499, 2021, doi: 10.1093/comjnl/bxaa172.

[4] Z. Drus and H. Khalid, "Sentiment analysis in social media and its application: Systematic literature review," *Procedia Comput. Sci.*, vol. 161, pp. 707–714, 2019, doi: 10.1016/j.procs.2019.11.174.

[5] M. Guia, R. R. Silva, and J. Bernardino, "Comparison of Naive Bayes, support vector machine, decision trees and random forest on sentiment analysis," *IC3K 2019 - Proc. 11th Int. Jt. Conf. Knowl. Discov. Knowl. Eng. Knowl. Manag.*, vol. 1, no. Ic3k, pp. 525–531, 2019, doi: 10.5220/0008364105250531.

[6] I. Technology, F. S. Nufus, W. Gata, C. Science, and U. N. Mandiri, "SENTIMENT ANALYSIS OF PUBLIC OPINION ON TRANSPORTATION," vol. 22, no. 2, 2025.

[7] S. Mutmainah and E. E. Citra, "Improving the Accuracy of Social Media Sentiment Classification with the Combination of TF-IDF Method and Random Forest Algorithm Peningkatan Akurasi Klasifikasi Sentimen Media Sosial dengan Kombinasi Metode TF-IDF dan Algoritma Random Forest," vol. 1, no. 1, pp. 30–40, 2025.

[8] M. A. Khder, "Web scraping or web crawling: State of art, techniques, approaches and application," *Int. J. Adv. Soft Comput. its Appl.*, vol. 13, no. 3, pp. 144–168, 2021, doi: 10.15849/ijasca.211128.11.

[9] D. Van Huan, N. T. B. Tuyet, T. M. Dung, V. T. Ha, and T. D. Duc, "The influence of Facebook intensity on life satisfaction and academic performance of college students during COVID-19 pandemic," *J. Affect. Disord. Reports*, vol. 16, no. April, p. 100782, 2024, doi: 10.1016/j.jadr.2024.100782.

[10] Y. HaCohen-Kerner, D. Miller, and Y. Yigal, "The influence of preprocessing on text classification using a bag-of-words representation," *PLoS One*, vol. 15, no. 5, pp. 1–22, 2020, doi: 10.1371/journal.pone.0232525.

[11] S. Vijayarani, M. J. Ilamathi, and M. Nithya, "Preprocessing Techniques for Text Mining - An Overview," vol. 5, no. 1, pp. 7–16.

[12] A. I. Kadhim, "An Evaluation of Preprocessing Techniques for Text Classification," vol. 16, no. 6, pp. 22–32, 2018.

[13] F. Method, B. On, T. Of, and I. Citizens, "Sentiment Analysis Terkait Pemindahan Ibu Kota Indonesia Menggunakan Metode Random Forest Berdasarkan Tweet Warga Negara Indonesia," vol. 3, no. 2, pp. 154–165, 2022.

[14] D. W. Oard and G. Marchionini, "A Conceptual Framework for Text Filtering," pp. 1–32.

[15] D. E. Birba, "study of data splitting algorithms for machine learning," 2020.

[16] W. N. I. Al-obaydy, H. A. Hashim, Y. A. Najm, and A. A. Jalal, "Document classification using term frequency-inverse document frequency and K-means clustering," vol. 27, no. 3, pp. 1517–1524, 2022, doi: 10.11591/ijeecs.v27.i3.pp1517-1524.

[17] H. Zhou, "Research of Text Classification Based on TF-IDF and CNN-LSTM," *J. Phys. Conf. Ser.*, vol. 2171, no. 1, 2022, doi: 10.1088/1742-6596/2171/1/012021.

[18] H. Liu, X. Chen, and X. Liu, "A Study of the Application of Weight Distributing Method Combining Sentiment Dictionary and TF-IDF for Text Sentiment Analysis," *IEEE Access*, vol. 10, pp. 32280–32289, 2022, doi: 10.1109/ACCESS.2022.3160172.

[19] S. Qaiser and R. Ali, "Text Mining : Use of TF-IDF to Examine the Relevance of Words to Documents," vol. 181, no. 1, pp. 25–29, 2018.

[20] G. Salton and C. Buckley, "TEXT RETRIEVAL," vol. 24, no. 5, pp. 513–523, 1988.

[21] P. Karthika, R. Murugeswari, and R. Manoranjithem, "Sentiment Analysis of Social Media Network Using Random Forest Algorithm," *IEEE Int. Conf. Intell. Tech. Control. Optim. Signal Process. INCOS 2019*, pp. 1–5, 2019, doi: 10.1109/INCOS45849.2019.8951367.

[22] H. A. Salman, A. Kalakech, and A. Steiti, "Random Forest Algorithm Overview," vol. 2024, pp. 69–79, 2024.

[23] A. Ahmed Bilal, O. Ayhan Erdem, and S. Toklu, "Children's Sentiment Analysis from Texts by Using Weight Updated Tuned with Random Forest Classification," *IEEE Access*, vol. 12, no. May, pp. 70089–70104, 2024, doi: 10.1109/ACCESS.2024.3400992.

[24] N. A. Semary, W. Ahmed, K. Amin, P. Pławiak, and M. Hammad, "Enhancing machine learning-based sentiment analysis through feature extraction techniques," *PLoS One*, vol. 19, no. 2 February, 2024, doi: 10.1371/journal.pone.0294968.

[25] C. A. Nurhaliza Agustina, R. Novita, Mustakim, and N. E. Rozanda, "The Implementation of TF-IDF and Word2Vec on Booster Vaccine Sentiment Analysis Using Support Vector Machine Algorithm," *Procedia Comput. Sci.*, vol. 234, pp. 156–163, 2024, doi: 10.1016/j.procs.2024.02.162.

[26] V. A. Fitri, R. Andreswari, and M. A. Hasibuan, "Sentiment analysis of social media Twitter with case of Anti-LGBT campaign in Indonesia using Naïve Bayes,

decision tree, and random forest algorithm," *Procedia Comput. Sci.*, vol. 161, pp. 765–772, 2019, doi: 10.1016/j.procs.2019.11.181.

[27] Z. Xiao, L. Wang, and J. Y. Du, "Improving the performance of sentiment classification on imbalanced datasets with transfer learning," *IEEE Access*, vol. 7, pp. 28281–28290, 2019, doi: 10.1109/ACCESS.2019.2892094.

[28] R. Obiedat *et al.*, "Sentiment Analysis of Customers' Reviews Using a Hybrid Evolutionary SVM-Based Approach in an Imbalanced Data Distribution," *IEEE Access*, vol. 10, pp. 22260–22273, 2022, doi: 10.1109/ACCESS.2022.3149482.

[29] Q. A. Xu, V. Chang, and C. Jayne, "A systematic review of social media-based sentiment analysis: Emerging trends and challenges," *Decis. Anal. J.*, vol. 3, no. May, p. 100073, 2022, doi: 10.1016/j.dajour.2022.100073.

[30] D. Tiwari, B. Nagpal, B. S. Bhati, A. Mishra, and M. Kumar, *A systematic review of social network sentiment analysis with comparative study of ensemble-based techniques*, vol. 56, no. 11. Springer Netherlands, 2023. doi: 10.1007/s10462-023-10472-w.