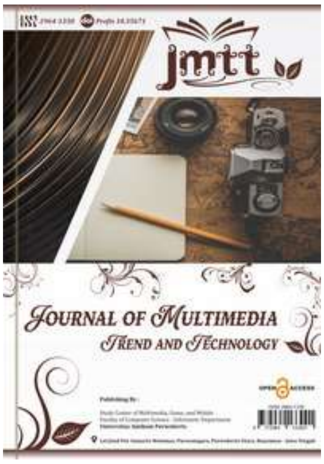
Journal of Multimedia Trend and Technology - JMTT
Volume 5, No. 1, April 2026, ISSN 2964-1330
https://journal.educollabs.org/index.php/jmtt/

Sentiment Analysis in User Reviews of Tourist Attractions in East Nusa Tenggara Using Machine Learning Classification

Aulia Dian Agustina [1*], Primandani Arsi [2], Pungkas Subarkah [3], Irfan Santiko [4]

[1,2,3,4] Informatic Departement, Faculty of Computer Science, Universitas Amikom Purwokerto, Purwokerto, Indonesia

ARTICLE INFO



History :

Submit on 5 January 2026
Review on 27 January 2026
Accepted on 2 March 2026

Keyword :

Sentiment Analysis;
Support Vector Machine;
TF-IDF;
Google Maps;
Travel Reviews

ABSTRACT

This study aims to analyze user review sentiments for six tourist attractions in East Nusa Tenggara Province by utilizing a large amount of review data obtained from Google Maps. Data was collected through a scraping process using Serp API, followed by cleaning and text pre-processing to improve data quality. Sentiment labeling was performed automatically using the Indo-BERT model to obtain three sentiment classes: positive, negative, and neutral. Text feature representation was performed using the Term Frequency–Inverse Document Frequency (TF-IDF) method, then classified using the baseline Support Vector Machine (SVM) model and the optimized SVM model with Grid-Search CV. The evaluation results showed that the baseline SVM model produced an accuracy of 83.87%, but showed an imbalance in performance between classes with a Macro F1-score of 0.4287. After parameter optimization using Grid-Search CV, the optimized SVM model produced an accuracy of 78.27% with an increase in the Macro F1-score value to 0.4818. This increase indicates an improvement in the model's ability to recognize minority sentiment classes despite a decrease in overall accuracy. Overall, the optimized SVM model provides more balanced and representative classification results in describing tourists' perceptions based on online reviews, so it can be used as a basis for sentiment analysis in the tourism sector.

***Corresponding Author:**

Aulia Dian Agustina
Informatic Departement, Faculty of Computer Science, Universitas Amikom Purwokerto, Jl Letjen Pol. Sumarto, Purwanegara, North Purwokerto, Indonesia
Email: auliadian@gmail.com

INTRODUCTION

The development of information technology and digital platforms has transformed the way people seek and share information, including in the tourism sector [1]. Today, prospective tourists no longer rely solely on official promotions or recommendations from specific sources, but also consider the experiences of other travelers shared through online reviews [2]. These reviews are an important source of information because they provide direct assessments of tourist destinations based on real-life visitor experiences [3][4].

East Nusa Tenggara (NTT) Province is one of the regions in Indonesia with significant tourism potential, both in terms of natural beauty and cultural richness. As tourism activity increases, various tourist attractions in NTT receive numerous reviews from visitors through digital platforms such as Google Maps. These reviews reflect perceptions, satisfaction, and complaints about the facilities and services available at these attractions [5][6].

Unlike conventional survey data, online reviews are unstructured, freely written, and use a diverse language style [7]. Furthermore, the number of available reviews is relatively large and continues to grow. This makes manual analysis of reviews inefficient and potentially subjectivist. Therefore, an approach capable of automatically and systematically processing review text data is needed so that the information contained therein can be optimally utilized [8].

Sentiment analysis is one approach that can be used to understand user opinion trends based on review text [9][10]. Through sentiment analysis, visitor reviews can be grouped into positive, negative, or neutral sentiment categories [11]. This grouping can provide an overview of visitor perceptions of a tourist attraction and help identify aspects that are considered good and those that need improvement [12][13].

In its application, sentiment analysis of review data requires a classification method capable of handling high-dimensional text data and quite complex language variations [14]. Machine learning methods are widely used in sentiment analysis due to their ability to recognize patterns in text data [15][16]. One algorithm frequently used for text classification is the Support Vector Machine (SVM), which is known to perform well on text data and is capable of producing stable results on small to medium-sized datasets [17].

Based on this background, this study was conducted to analyze the sentiment of visitor reviews of tourist attractions in East Nusa Tenggara using a machine learning approach. By utilizing online reviews as a data source and applying appropriate classification methods, this study is expected to provide a more objective picture of visitor perceptions of tourist attractions in East Nusa Tenggara (NTT). The results are expected to serve as evaluation material and considerations for relevant parties in the development and management of tourist attractions in East Nusa Tenggara. This study aims to analyze and prove the ability of the Support Vector Machine (SVM) method in processing and classifying sentiments in large amounts of review data (big data) obtained from the Google Maps platform, as well as evaluating the model's performance in conditions of unbalanced data distribution.

In general, sentiment analysis can be performed at several levels: document-level, sentence-level, and aspect-level. This study uses a document-level sentiment analysis approach, where each review is treated as a single document representing a single sentiment tendency. This approach is considered appropriate for analyzing travel reviews because most visitor comments are written in short paragraphs that, overall, reflect the primary impression of the visited tourist attraction [18].

However, sentiment analysis in travel reviews presents its own challenges. Review texts are generally short, use informal language, and contain variations in writing, emoticons, and informal language. Furthermore, the distribution of sentiment in travel reviews is often unbalanced, with positive sentiment tending to outweigh negative sentiment. This condition can impact the performance of classification models if not handled properly, requiring preprocessing and selecting appropriate methods to ensure accurate and reliable analysis results.

In this study, sentiment analysis was used as the basis for grouping visitor reviews of tourist attractions in East Nusa Tenggara into specific sentiment classes. The results of this grouping are expected to provide a general overview of visitor perceptions of the tourist attractions being analyzed, as well as being the basis for applying the classification method using the Support Vector Machine (SVM) algorithm in the next stage [19].

In the context of this research, the term "big data" refers to large amounts of unstructured review data that require computational processing for analysis. The dataset used consisted of 3,049 text reviews from Google Maps, which are practically impossible to analyze manually. Therefore, a machine learning approach was used to systematically process and evaluate this data [20].

METHOD

This research was conducted online, utilizing user review data available on the Google Maps platform. The research focused on several tourist destinations in East Nusa Tenggara Province, namely Gili Lawa Darat, Kelor Island, Manjarite Beach, Pink Beach, Padar Island, and Puncak Waringin. These destinations were selected based on the availability of adequate review data and the level of user activity in providing reviews, thus ensuring the data obtained was representative enough for analysis in this study.

The user review data collection process (web scraping) was conducted on Saturday, July 26, 2025, and continued until Monday, January 12, 2026. After data collection was completed, the research continued with data cleaning, text pre-processing, sentiment labeling, text feature extraction, classification model training using the Support Vector Machine (SVM) algorithm, and parameter optimization (hyperparameter tuning) to achieve optimal model performance. The final stage of the research included evaluation and interpretation of the sentiment classification results based on predetermined evaluation metrics. All research activities were carried out in stages and continuously in accordance with the established research objectives.

The research concepts and stages used are as follows:

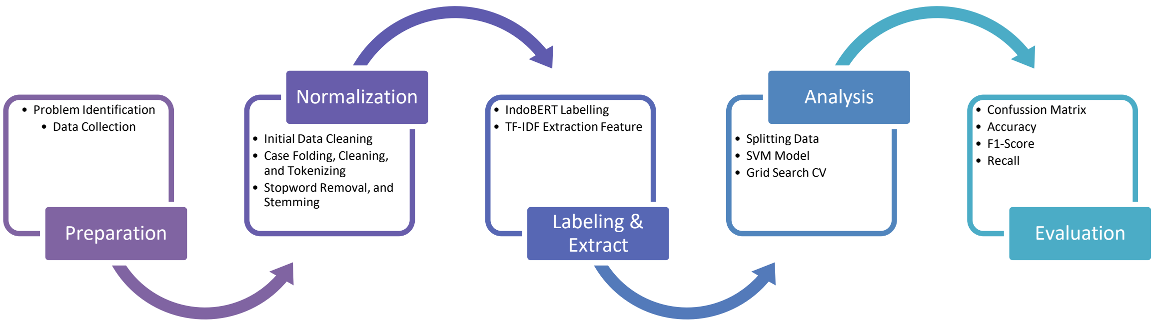


Figure 1. Research concepts and stages

The data collection phase is the initial stage in this research, aiming to obtain a user review dataset as the primary source for the sentiment analysis process. The data used consisted of user reviews on the Google Maps platform for six tourist destinations in East Nusa Tenggara Province: Gili Lawa Darat, Kelor Island, Manjarite Beach, Pink Beach, Padar Island, and Puncak Waringin. Google Maps was selected as the data source based on the platform's characteristics, which provide reviews based on direct user experience in text form, making it relevant for describing tourist perceptions of tourist attractions.

At this stage, several data cleaning processes were performed, including removing reviews without comment text, eliminating duplicate data, and filtering out reviews irrelevant to the research objectives. Reviews containing only symbols, random characters, or information that does not reflect user opinions were also removed from the dataset. This process was carried out to reduce noise that could affect the sentiment analysis results.

The text pre-processing phase is a crucial stage in this research, aiming to prepare the text data for optimal processing in the sentiment analysis phase. User review data obtained from Google Maps still contains various spelling variations, symbols, and irrelevant words. Therefore, pre-processing is necessary to improve data quality before analysis using machine learning methods.

The sentiment labeling stage is carried out to assign a sentiment class label to each user review that has undergone text pre-processing. Review data obtained from Google Maps does not have a direct sentiment label, so a labeling process is necessary so that the data can be used in training a sentiment classification model. In this study, sentiment labeling was performed using the IndoBERT model. IndoBERT is used as a tool to automatically and consistently predict review sentiment tendencies. Each review is classified into one of three sentiment classes: positive, negative, and neutral, based on the model's predictions.

Text feature extraction is a crucial stage in this study, aiming to convert the pre-processed text data into numerical form so that it can be processed by the classification algorithm.

In this study, the feature extraction method used is Term Frequency–Inverse Document Frequency (TF-IDF). This method was chosen because it is able to give weight to each word based on its level of importance in a document or in the entire collection of documents. Term Frequency (TF) is used to measure how frequently a word appears in a document. The more frequently a word appears in a review, the higher its TF value. However, the TF value only considers the frequency of a word's appearance in a single document without considering its distribution across other documents. Mathematically, Term Frequency is formulated as follows (1):

$$\mathrm{TF}(t,d) = \frac{f(t,d)}{\sum_{t^1 \in d} f(t^1,d)} \quad (1)$$

Description:
f(t,d) = number of occurrences of term t in document d
$\sum_{t^1 \in d} f(f^1, d)$ = total of all terms in the document

Inverse Document Frequency (IDF) is used to reduce the weight of words that frequently appear in many documents, as these words are considered less informative in distinguishing one document from another. Conversely, words that rarely appear in a collection of documents will have a higher IDF value. The formula for Inverse Document Frequency is as follows (2):

$$\mathrm{IDF}(t) = \log\left(\frac{N}{df(t)}\right) \quad (2)$$

Description:
N = total number of documents in the dataset
df(t) = number of documents containing term t

The final TF-IDF score is obtained by multiplying the Term Frequency (TF) and Inverse Document Frequency (IDF) values. This score represents the importance of a word in a document relative to the entire document. Mathematically, TF-IDF is formulated as follows (3):

$$\mathrm{TF-IDF}(t,d) = \mathrm{TF}(t,d) \times \mathrm{IDF}(t) \tag{3}$$

The result of the feature extraction process using the TF-IDF method is a numeric vector representing each user review. This feature vector is then used as input to train a sentiment classification model using the Support Vector Machine (SVM) algorithm.

RESULT

Data Collecting.

In this study, the data was split into 80% training data and 20% testing data. The training data was used to train the classification model using the Support Vector Machine (SVM) algorithm, while the testing data was used to evaluate the model's performance based on sentiment prediction results. This split ratio was chosen because it provided a balance between sufficient data for model training and representative data for testing.

In this study, no data balancing was performed on the sentiment classes. The distribution of sentiment classes was left to reflect the original user review data obtained from Google Maps. This decision was made to ensure that the classification results reflect the actual user perceptions of the tourist attraction under study and to avoid potential biases that could arise from artificial data balancing. The separated training and testing data were then used in the next stage: classification model training and model performance evaluation.

The sentiment classification stage in this study was conducted using the Support Vector Machine (SVM) algorithm. SVM is a machine learning algorithm widely used in sentiment analysis due to its excellent ability to handle high-dimensional data, such as text data extracted from TF-IDF features. The working principle of SVM is to find an optimal hyperplane that can separate data into sentiment classes with maximum margin. The optimal hyperplane is chosen in such a way that the distance between the hyperplane and the closest data points from each class (Support vectors) is maximized, thus ensuring the model has good generalization capabilities to new data.

Based on the initial data collection results, a total of 3,897 reviews were obtained. A summary of the number of reviews for each tourist attraction is presented in Table 1.

Table 1. Number of Initial Reviews per Attraction

No	Object	Review Count
1	Gili Lawa Darat	82
2	Kelor Island	338
3	Manjarite Beach	354
4	Pink Beach	1038
5	Pulau Padar	1038
6	Puncak Waringin	1047
	Total	3897

Table 1 shows that the number of reviews received for each tourist attraction is uneven. Puncak Waringin is the tourist attraction with the most reviews, with 1,047 reviews, followed by Padar Island and Pink Beach, with 1,038 reviews each. The high number of reviews for these three attractions indicates a relatively high level of tourist visits and high user interaction on the Google Maps platform. Kelor Island and Manjarite Beach have a lower number of reviews, with 338 reviews, and Gili Lawa Darat has the fewest reviews, with 82 reviews. This difference in the number of reviews may be influenced by several factors, such as limited access to the location, the popularity of the tourist destination, differences in tourist visit volume, and variations in user participation in providing online reviews. To provide a visual representation of the distribution of reviews for each tourist attraction, the collected data is also presented in bar graph form in Figure 2.

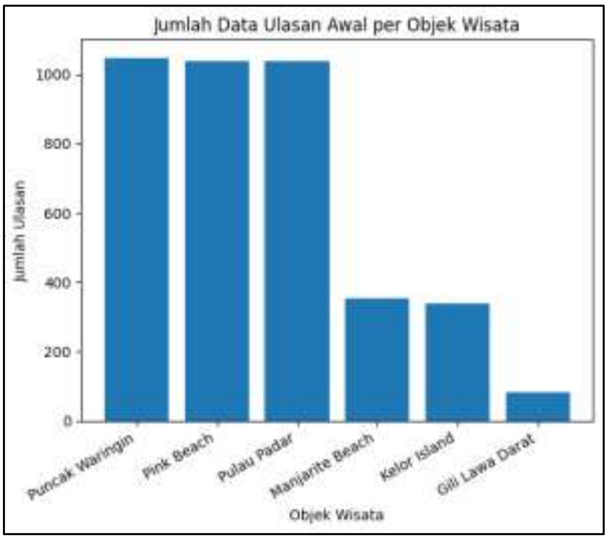


Figure 2. Initial Number of Tourist Attraction Reviews Graph

Cleaning Data.

The data cleaning phase was conducted after all review data was successfully collected from the Google Maps platform through a scraping process. The data obtained in the initial phase amounted to 3,897 reviews and was still raw, making it not yet fully ready for sentiment analysis. The initial dataset still contained empty reviews, duplicate data, and text that was too short and did not clearly represent opinions. The cleaning process was carried out systematically through several stages, namely:

a. Standardization and selection of relevant attributes, including the tourist_object, name, rating, date, review, and local_guide columns.
b. Removal of empty or null reviews.
c. Removal of duplicate data based on content similarity in the review columns.
d. Filtering of reviews with text lengths less than two words.

After all cleaning stages were implemented, the data volume was reduced to 2,942 reviews, resulting in 955 reviews being removed at this stage. A summary of the cleaning results is presented in Table 2.

Table 2. Data Cleaning Results Summary

No	Description	Value
1	Initial data count (raw)	3.897
2	Data count after cleaning	2.942
3	Data count deleted	955
4	Initial column count	6
5	Column count after cleaning	6

Recapitulation of the number of reviews per tourist attraction after the Cleaning process was carried out to ensure that data deletion did not significantly change the representation of data distribution.

Table 3. Number of Reviews per Attraction After Cleaning

No	Destination Object	Review Count (after cleaning)
1	Pulau Padar	988
2	Pink Beach	925
3	Puncak Waringin	581
4	Kelor Island	220
5	Manjarite Beach	169
6	Gili Lawa Darat	59
	Total	2.942

The distribution of reviews after the cleaning process shows that the data distribution pattern between tourist attractions remains consistent with the initial conditions. Padar Island and Pink Beach still have the highest number of reviews, while Gili Lawa Darat has the lowest number of reviews. This indicates that the cleaning process successfully improved the quality of the dataset without significantly changing the characteristics of the data distribution. With the completion of the cleaning stage, the dataset is considered to have met data quality criteria and is ready for use in the text pre-processing stage and sentiment analysis using the Support Vector Machine (SVM) method.

Label and Extraction Data.

The sentiment labeling stage is carried out to assign sentiment classes to review data that has undergone cleaning and text preprocessing. This labeling is necessary because review data obtained from the Google Maps platform does not have a direct sentiment label. With sentiment labels, each review can be used as labeled data in the sentiment classification process using the Support Vector Machine (SVM) algorithm.

Sentiment labeling in this study was performed automatically using the transformer-based IndoBERT model, w11wo/indonesian-roberta-base-sentiment-classifier. This model was chosen because it was trained on an Indonesian language corpus and is capable of classifying text sentiment into three main classes: positive, negative, and neutral.

The text used as input for the labeling process was the review column (the original text), because the transformer-based model performs best on text that retains the sentence structure and context intact. The sentiment labeling results yielded a class distribution of 2,942 valid reviews after the cleaning and preprocessing processes. A summary of the number of reviews in each sentiment class is presented in Table 4.

Table 4. Distribution of Sentiment Labeling Results Using IndoBERT

Sentimen	Jumlah
Positif	2.436
Negatif	349
Netral	157
Total	2.942

Based on Table 4, positive sentiment dominates the dataset, with 2,436 reviews. Negative sentiment accounts for 349 reviews, while neutral sentiment accounts for 157. Proportionally, positive sentiment comprises approximately 82.8% of the total data, while negative and neutral sentiments account for 11.9% and 5.3%, respectively. The dominance of positive sentiment indicates that the majority of users provided appreciative reviews of tourist attractions in East Nusa Tenggara Province. However, this uneven distribution also indicates an imbalanced dataset, with positive sentiment being the dominant class compared to negative and neutral.

To clarify the comparison of data volumes within each sentiment class, the labeling results are visualized in the form of a bar graph, as shown in Figure 3.

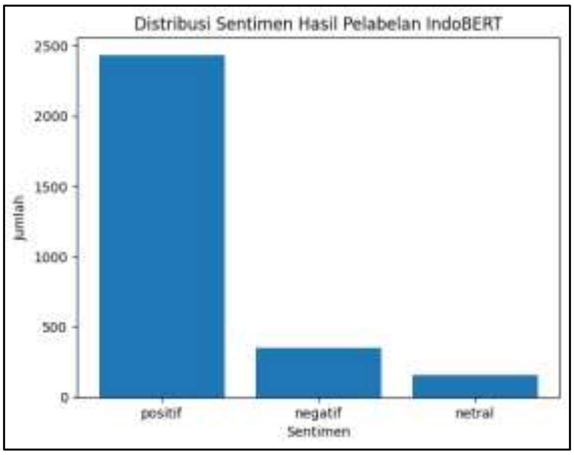


Figure 3. Distribution of User Review Sentiment from IndoBERT Labeling Results

The TF-IDF method calculates the weight of each word based on two main components: the frequency of the word's appearance in a single review document (term frequency) and the word's rarity across all documents (inverse document frequency). Words that appear frequently in certain reviews, but rarely in others, will have a higher TF-IDF weight. This condition indicates that the word is more representative in differentiating sentiment characteristics between reviews. To illustrate the results of TF-IDF feature extraction, Table 5 presents the ten words with the highest TF-IDF weights obtained from user review data. The table shows that each word has a different weight, depending on its level of importance in the overall dataset.

Table 5. Example of TF-IDF Feature Representation on Review Data

No	Word Feature	Value *TF-IDF*
1	indah	0,038968
2	the	0,037306
3	bagus	0,028601
4	view	0,023785
5	pantai	0,023347
6	pandang	0,023012
7	pink	0,020215
8	pulau	0,019499
9	to	0,019273
10	beach	0,018538

Words with high TF-IDF weights, such as "indah," "lihat," and "baik," indicate that they frequently appear in specific reviews, but do not dominate the entire document. This makes them more discriminating in representing review sentiment. Conversely, words with lower TF-IDF weights tend to appear evenly across many reviews, making their contribution to distinguishing sentiment relatively smaller.

To complement the analysis in the feature extraction stage, a word cloud visualization of dominant words was performed. This visualization aims to provide an exploratory overview of the most frequently occurring words in the review data after text cleaning and pre-processing. The wordcloud is built from the clean_reviews column, so the words displayed are free of noise such as punctuation, numbers, and meaningless words. The word size in the wordcloud visualization represents the frequency of word occurrences in the entire review dataset.

Figure 4. Wordcloud of Dominant Words in Travel Reviews

Classification Results Using Support Vector Machine (SVM) Baseline.

The data used in the classification stage consists of review data that has undergone all initial processing stages, including cleaning, text preprocessing, and automatic sentiment labeling using the IndoBERT model. All these stages were carried out to ensure that the data used was of adequate quality and met the requirements of sentiment classification analysis. The final dataset used in the classification process consisted of 2,942 reviews, each with a sentiment label: positive, negative, or neutral. These sentiment labels make the dataset labeled data, allowing it to be used directly in the training and testing of sentiment classification models using the Support Vector Machine (SVM) algorithm. The preprocessed review text was then represented numerically using the Term Frequency Inverse Document Frequency (TF-IDF) method. This numerical representation is necessary because the SVM algorithm cannot process text data directly. Using TF-IDF, each review was converted into a 5,000-feature numeric vector, according to the max_features=5000 parameter in the feature extraction process. The agreement between the review data, IndoBERT-labeled sentiment labels, and TF-IDF numeric features indicates that all data components are completely and consistently structured. The dataset was then divided into training and test data with an 80:20 ratio using the stratified split method, thus maintaining the distribution of sentiment classes in both data subsets. The training data consisted of 2,353 reviews, while the test data consisted of 589 reviews. With this data ready, the baseline SVM model development process can be carried out to evaluate the initial performance of sentiment classification before parameter optimization.

The sentiment labeling dataset, consisting of 2,942 reviews, was divided into training and test data using a train-test split method with a proportion of 80% for training data and 20% for testing data. The division was performed using stratified sampling parameters to maintain the distribution of sentiment classes in both data subsets. The results show that the training data consists of 2,353 reviews, while the test data consists of 589 reviews. The distribution of sentiment labels in the training data is as follows:

Table 6. Sentiment Distribution on Training Data (80%)

Sentiment	Count
Positif	1948
Negatif	279
Netral	126
Total	2.353

Table 7. Sentiment Distribution on Test Data (20%)

Sentiment	Count
Positif	488
Negatif	70
Netral	31
Total	589

Further evaluation was conducted using a classification report to measure model performance for each sentiment class based on precision, recall, and F1-score metrics. A summary of the baseline SVM model classification report results is shown in Table 8.

Table 8. Summary of Classification Report Results of Baseline SVM Model

Sentiment	Precision	Recall	F1-Score	*Support*
Negatif	0,5714	0,1714	0,2637	70
Netral	0,3333	0,0645	0,1081	31
Positif	0,8541	0,9836	0,9143	488
Macro Average	0,5863	0,4065	0,4287	589
Weighted Average	0,7931	0,8387	0,7945	589

Based on the Classification Report results of the baseline SVM model, the positive sentiment class performed best with a precision value of 0.8541, a recall of 0.9836, and an F1-score of 0.9143. The very high recall value indicates that almost all positive data was successfully recognized by the model. Conversely, the model's performance in the negative and neutral sentiment classes was still low. In the negative class, the recall was only 0.1714, while in the neutral class it was 0.0645. This indicates that most data in the minority class was still misclassified as positive sentiment. The macro F1-score value of 0.4287 reflects that the model's performance was not yet balanced between classes, which was influenced by the dominance of the positive class in the dataset.

Visual evaluation of the model prediction results was carried out using a confusion matrix, as shown in Figure 5.

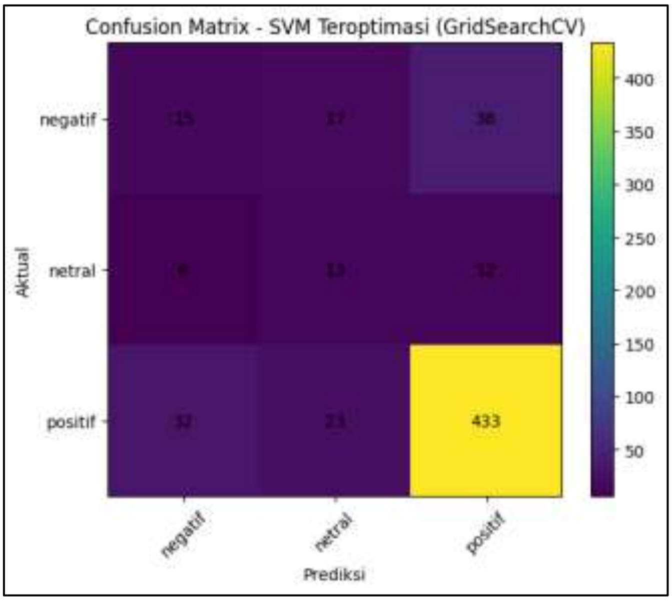


Figure 5. Optimized Visual Confusion Matrix SVM (GridSearchCV)

Based on the confusion matrix, the following details were obtained:

a. Of the 70 test data sets labeled negative, 12 were correctly predicted as negative, 3 were predicted as neutral, and 55 were predicted as positive.
b. Of the 31 test data sets labeled neutral, 2 were correctly predicted as neutral, 2 were predicted as negative, and 27 were predicted as positive.
c. Of the 488 test data sets labeled positive, 480 were correctly predicted as positive, while 7 were predicted as negative and 1 was predicted as neutral.

These findings indicate that model performance is not evenly distributed across classes. The dominance of the positive class in the dataset causes the model to predict this class more frequently, thus limiting its ability to recognize variations in minority sentiment.

Based on the evaluation results, the baseline SVM model achieved an accuracy of 83.87% on the test data. This value indicates that the model is generally capable of correctly classifying most of the data. However, further analysis using the classification report and confusion matrix revealed that the model's performance was not uniform across all sentiment classes.

The model's dominant ability in recognizing the positive sentiment class was achieved with a recall value of 0.9836, indicating that almost all reviews labeled positive were correctly predicted. Conversely, its performance in the negative and neutral sentiment classes was relatively low, with recall values of 0.1714 and 0.0645, respectively. This indicates that most data labeled negative and neutral still tended to be predicted as positive.

This imbalance in performance between classes is reflected in the Macro F1-score of 0.4287, which is significantly lower than the accuracy value. This difference indicates that, despite high accuracy, the model's ability to handle minority classes is not yet balanced. This condition is influenced by the imbalanced distribution of the dataset, where the positive sentiment class dominates approximately 82.8% of the total test data. The dominance of the majority class causes the model to form a decision boundary that favors the positive class to maximize overall accuracy.

These results also indicate that, based on the labeling process, the majority of user reviews of tourist attractions in East Nusa Tenggara Province are categorized as positive sentiment. However, the presence of negative and neutral sentiment remains important in the analysis, as these two classes potentially contain information about aspects of services, facilities, and tourism experiences that require further evaluation.

Thus, although the baseline SVM model did not demonstrate uniform performance across all classes, the classification results provided an initial overview of the distribution of visitor sentiment. To improve the balance of classification performance between classes, a parameter optimization process using the GridSearchCV method is required, which is discussed in the next section.

To provide a more comprehensive overview of the performance differences between the baseline Support Vector Machine (SVM) model and the optimized SVM model from GridSearchCV, a comparison was conducted based on the main evaluation metrics, namely accuracy, precision, recall, F1-score, and macro F1-score. A summary of the evaluation results of both models is presented in Table 9.

Table 9. Summary of SVM Model Performance Comparison

Metrix Evaluasi	***SVM Baseline***	***Optimization SVM***
Akurasi	0,8387	0,7827
Precision (Negatif)	0,5714	0,2830
Recall (Negatif)	0,1714	0,2143
F1-Score (Negatif)	0,2637	0,2439
Precision (Netral)	0,3333	0,2453
Recall (Netral)	0,0645	0,4194
F1-Score (Netral)	0,1081	0,3095
Macro F1-Score	0,4287	0,4818

Based on Table 4.15, the baseline SVM model has a higher accuracy value, at 0.8387, compared to the optimized SVM model's 0.7827. This decrease in accuracy occurs because the optimized model no longer focuses so much on the majority class (positive sentiment), but instead attempts to improve detection capabilities for the minority classes, namely negative and neutral. The most significant improvement is seen in the neutral sentiment class. The recall value for the neutral class increased significantly from 0.0645 in the baseline model to 0.4194 in the optimized model. This indicates that after parameter optimization, the model is much better able to recognize neutral data that was previously frequently misclassified.

In the negative sentiment class, the recall value also increased from 0.1714 to 0.2143, although the increase was not as significant as for the neutral class. However, the precision value for the negative class decreased, indicating an increase in negative predictions, but not yet fully accurate. Overall, the improvement in model quality was most evident in the Macro F1-score, which increased from 0.4287 to 0.4818. The Macro F1-score was used because it reflects the balance of performance between classes without being influenced by the dominance of the majority class. This increase indicates that the optimized SVM model performed more evenly across all sentiment classes compared to the baseline model. Thus, despite the decrease in accuracy, the optimized SVM model demonstrated an improvement in its ability to recognize minority classes and produced more balanced classification performance.

CONCLUTIONS

Based on the results of a study of 610 Google Maps user review datasets for six tourist attractions in East Nusa Tenggara Province, the following conclusions can be drawn:

1. The data collection, cleaning, and text pre-processing processes successfully produced a dataset suitable for use in machine learning-based sentiment analysis.
2. Labeling results using the IndoBERT model indicate that the majority of reviews have positive sentiment, reflecting tourists' generally favorable perceptions of the tourist attractions analyzed.
3. The baseline SVM model achieved an accuracy of 83.87%, but showed uneven performance in the minority sentiment class. This is evident from the negative recall value of 0.1714 and the neutral recall of 0.0645, indicating that the model is still heavily influenced by the dominance of the positive class.
4. After parameter optimization using GridSearchCV, the optimized SVM model achieved an accuracy of 78.27%, with an increase in the Macro F1 value from 0.4287 to 0.4818, indicating improved performance balance between classes.
5. The optimized model demonstrated improved ability to recognize minority classes, particularly the neutral class, with recall increasing from 0.0645 to 0.4194, and recall increasing for the negative class from 0.1714 to 0.2143.
6. Overall, the optimized SVM model produced a classification that was more representative of the sentiment distribution in the dataset, as it no longer focused too heavily on the majority class. This model is more suitable for use as a basis for analyzing tourist perceptions based on online reviews.

The sentiment analysis results indicate that while the majority of reviews are positive, there are still negative and neutral sentiments that require attention. Tourist attraction managers are advised to use the sentiment analysis results as a basis for regular evaluations of service quality, facilities, cleanliness, accessibility, and the overall visitor experience. Negative sentiment can be used as a basis for service improvements, while positive sentiment can be utilized as a promotional strategy and to strengthen the destination's image.

Acknowlegdement

The author would like to express his gratitude to Mrs. Primandani Arsi, S.ST., M.Kom. and Mr. Pungkas Subarkah, M.Kom., as the research promoter lecturers, who have patiently guided, directed, and provided very meaningful input during the process of compiling this thesis. Their sincerity, discipline, and attention have become valuable provisions for the author in completing this thesis as well as possible. Mr. Dr. Irfan Santiko, S.Kom., M.Kom. as the Academic Promoter Lecturer, who has provided direction, motivation, and support from the beginning of the author's academic journey until the completion of the study.

Author Contributions

A.D.A, P.A. Conceptual Idea, A.D.A, Writing and Collecting Data, P.S, I.S, Corrector Analysis.

Funding

Not applicable.

REFERENCE

[1] H. Fang, G. Xu, Y. Long, and W. Tang, "An effective ELECTRA-based pipeline for sentiment analysis of tourist attraction reviews," *Appl. Sci.*, vol. 12, no. 21, p. 10881, 2022.
[2] F. Borrajo-Millán, M.-M. Alonso-Almeida, M. Escat-Cortes, and L. Yi, "Sentiment analysis to measure quality and build sustainability in tourism destinations," *Sustainability*, vol. 13, no. 11, p. 6015, 2021.
[3] Y. Qin, X. Wang, and Z. Xu, "Ranking tourist attractions through online reviews: A novel method with intuitionistic and hesitant fuzzy information based on sentiment analysis," *Int. J. fuzzy Syst.*, vol. 24, no. 2, pp. 755–777, 2022.
[4] I. P. Windasari and D. Eridani, "Sentiment analysis on travel destination in Indonesia," in *2017 4th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE)*, 2017, pp. 276–279.
[5] Y. Luo, X. Zhang, Y. Qin, Z. Yang, and Y. Liang, "Tourism Attraction Selection with Sentiment Analysis of Online Reviews Based on Probabilistic Linguistic Term Sets and the IDOCRIW-COCOSO Model: Y. Luo et al.," *Int. J. Fuzzy Syst.*, vol. 23, no. 1, pp. 295–308, 2021.
[6] M. A. Wirahudha, D. Damayanti, and D. A. Megawaty, "Sentiment Analysis of Comments on Tourist Attractions In Lampung Province Using The Naive Bayes Method," *J. Tek. Inform.*, vol. 6, no. 2, pp. 935–950, 2025.
[7] S. P. Saharani and W. Yustanti, "Sentiment Analysis and Topic Modeling of Tourist Attractions in Gresik Regency Using the BERT Method," *J. Emerg. Inf. Syst. Bus. Intell.*, vol. 7, no. 1, pp. 62–70, 2026.
[8] N. W. S. Saraswati, I. K. G. D. Putra, M. Sudarma, and I. M. Sukarsa, "The Image of Tourist Attraction in Bali Based on Big Data Analytics and Sentiment Analysis," in *2023 International Conference on Smart-Green Technology in Electrical and Information Systems (ICSGTEIS)*, 2023, pp. 82–87.
[9] D. I. Af'idah, R. Kusumaningrum, and B. Surarso, "Long short term memory convolutional neural network for Indonesian sentiment analysis towards touristic destination reviews," in *2020 International Seminar on Application for Technology of Information and Communication (iSemantic)*, 2020, pp. 630–637.
[10] A. Larasati, J. Sayono, A. Purnomo, E. Mohamad, M. Farhan, and P. Rahmawati, "Applying web mining and sentiment analysis to assess tourists review on Batu City tourist destination," in *2020 4th International Conference on Vocational Education and Training (ICOVET)*, 2020, pp. 63–68.
[11] E. P. Harahap, H. D. Purnomo, A. Iriani, I. Sembiring, and T. Nurtino, "Trends in sentiment of Twitter users towards Indonesian tourism: analysis with the k-nearest neighbor method," *Comput. Sci. Inf. Technol.*, vol. 5, no. 1, pp. 19–28, 2024.
[12] A. R. Setyawan, L. H. Suadaa, and B. Yuniarto, "Aspect-based sentiment analysis using adaptive aspect on tourist reviews in Jakarta," *Sist. J. Sist. Inf.*, vol. 13, no. 6, pp. 2456–2466, 2024.
[13] R. A. N. Nayoan, A. F. Hidayatullah, and D. H. Fudholi, "Convolutional neural networks for

Indonesian aspect-based sentiment analysis tourism review," in *2021 9th International Conference on Information and Communication Technology (ICoICT)*, 2021, pp. 60–65.

[14] E. D. Madyatmadja, J. Suhartono, C. Walandi, D. P. Hartanujaya, M. Budiharjo, and V. Febrian, "Harmonizing Sentiments: User Reviews from Social Media of Kuta Beach as a Tourist Attraction," in *2024 Beyond Technology Summit on Informatics International Conference (BTS-I2C)*, 2024, pp. 462–467.

[15] S. Kurniawan, A. S. Pramayoga, and Y. F. Ashari, "Sentiment Analysis in Coastal Tourism: A Systematic Literature Review," *Eng. Technol. \& Appl. Sci. Res.*, vol. 15, no. 6, pp. 30226–30233, 2025.

[16] I. Santiko, T. R. Soeprobowati, B. Surarso, and I. Tahyudin, "Traditional-Enhance-Mobile-Ubiquitous-Smart : Model Innovation in Higher Education Learning Style Classification Using Multidimensional and Machine Learning Methods," *J. Appl. Data Sci.*, vol. 6, no. 1, pp. 753–772, 2025, doi: https://doi.org/10.47738/jads.v6i1.598.

[17] A. Santosa, M. I. S. Kusnadi, A. H. H. Rumainum, and A. A. S. Gunawan, "Exploring Tourist Movement in Central Java Using Network Science and Sentiment Analysis," in *2025 4th International Conference on Electronics Representation and Algorithm (ICERA)*, 2025, pp. 310–315.

[18] D. D. Purwanto, "Empirical Evaluation of IndoBERT and LSTM for Sentiment Analysis of Tourism Reviews: A Data-Driven Study on Kenjeran Park," *J. Tek. Inform.*, vol. 7, no. 1, pp. 463–474, 2026.

[19] R. I. Perwira, V. A. Permadi, D. I. Purnamasari, R. P. Agusdin, and others, "Domain-specific fine-tuning of IndoBERT for aspect-based sentiment analysis in Indonesian travel user-generated content," *J. Inf. Syst. Eng. Bus. Intell.*, vol. 11, no. 1, pp. 30–40, 2025.

[20] J. Arnandy, I. Budi, A. Ramadiah, P. K. Putra, and A. B. Santoso, "Aspect-Based Sentiment Analysis In The Asean Tourism Sector: A Systematic Literature Review," *J. Impresi Indones.*, vol. 4, no. 11, pp. 5316–5328, 2025.