JTAM (Jurnal Teori dan Aplikasi Matematika)
http://journal.ummat.ac.id/index.php/jtam
p-ISSN 2597-7512 | e-ISSN 2614-1175
Vol. 10, No. 1, January 2026, pp. 219-229

Sentiment Analysis of Hotel Reviews in Senggigi using Decision Tree and Support Vector Machine Algorithm

Lalu Muhammad Reza Suganda Putra[1]*, Heri Wijayanto[1], I Gede Putu Wirarama Wedashwara[1]
[1]Department of Magister Information Technology, University of Mataram, Indonesia
rezasuganda89@gmail.com

Article History:
Received : 22-09-2025
Revised : 17-12-2025
Accepted : 31-12-2025
Online : 05-01-2026

Keywords:
Sentiment Analysis;
Hotel Reviews;
Senggigi;
Decision Tree;
Support Vector Machine.

ABSTRACT

The tourism industry is a rapidly growing sector that significantly contributes to the economy, including Indonesia. One of the popular tourist destinations in Indonesia is Senggigi, located on the island of Lombok. This destination offers high natural and cultural appeal. In the tourism industry, hotels are crucial as primary accommodations for travelers to stay and rest. Tourist reviews on hotel services greatly influence potential visitor's decisions in selecting the right accommodation. Therefore, sentiment analysis of hotel reviews is essential for understanding customer satisfaction levels and assisting hotel managers in improving service quality. This research applies a comparative quantitative approach using Decision Tree and Support Vector Machine (SVM) algorithms. The dataset consists of 6,920 hotel reviews collected from TripAdvisor platforms through web scraping techniques. Data preprocessing included data cleaning, case folding, tokenization, stop word removal, and stemming to enhance classification performance. Sentiment labels were categorized into positive, neutral, and negative classes. Model performance was evaluated using multiple metrics, including accuracy, precision, recall, and F1-score, to ensure a comprehensive assessment. The word frequency distribution reveals that that accommodation experience and room quality play a crucial role in customer satisfaction. Positive sentiment is dominated by adjectives like great, nice, and beautiful, reflecting pleasant experiences. Negative sentiment is expressed more politely through phrases such as not good or not very nice. Neutral sentiment tends to be descriptive without strong emotional expression. In terms of model performance, SVM outperformed the Decision Tree model, achieving an accuracy of 90%, precision of 91%, recall of 90%, and an F1-score of 85%. In comparison, the Decision Tree achieved an accuracy of 87%, precision of 84%, recall of 87%, and an F1-score of 85%. These findings demonstrate the superior capability of SVM in handling complex and diverse textual data. This study contributes academically by strengthening empirical evidence on the effectiveness of machine learning–based sentiment analysis in the tourism domain and practically by providing actionable insights for hotel managers to improve service quality and customer satisfaction.

https://doi.org/10.31764/jtam.v10i1.34960

———————— ◆ ————————

A. INTRODUCTION

The tourism industry has experienced rapid global growth and has become a vital contributor to economic development (Deepthi & Shariff, 2024). This growth is driven by increasing tourist mobility, digital transformation, and the expansion of tourism-related services. Tourism generates significant economic value by creating employment opportunities and supporting regional development (Hasibuan et al., 2023). As tourism demand continues to rise, competition among destinations and service providers intensifies, particularly in the hospitality sector, where service quality and customer satisfaction are essential for sustainability (Chi et al., 2025).

In Indonesia, tourism is recognized as one of the key pillars of national economic development. The country offers a wide variety of tourist destinations, supported by natural resources and cultural diversity, which attract both domestic and international visitors (Hasibuan et al., 2023). Along with increasing tourist arrivals, the hospitality industry has expanded rapidly, making hotels a critical component of the tourism value chain. The quality of hotel services plays a decisive role in shaping tourists' overall experiences and satisfaction levels (Chi et al., 2025; Ali et al., 2021), thereby influencing destination image and competitiveness.

West Nusa Tenggara (NTB) Province is well known for its pristine marine tourism, particularly around Lombok Island and the Gili Islands (Suhono et al., 2020). The natural beauty of beaches, marine ecosystems, and local culture has positioned the region as a prominent ecotourism destination (Islahuddin et al., 2024). One of the most popular destinations in Lombok is Senggigi Beach, which is renowned for its scenic landscape, tranquil atmosphere, and cultural richness (Wahyuningtyas et al., 2020). The increasing popularity of Senggigi Beach has encouraged hotel development in the area, intensifying competition and increasing the importance of understanding customer perceptions of hotel services.

In today's digital era, online reviews have become an essential source of information for tourists when choosing accommodation. Tourists frequently share their experiences through platforms such as TripAdvisor, Google Reviews, and social media, making user-generated content highly influential in shaping public perceptions of hotel quality (An & Ozturk, 2022; Burhanudin, 2024). Online hotel reviews reflect customers' real experiences regarding service quality, facilities, cleanliness, and staff performance. Consequently, analyzing these reviews is crucial for hotels to understand customer satisfaction and improve service performance (Chi et al., 2025).

Sentiment analysis is a widely used text mining technique for identifying opinions or emotions expressed in textual data and classifying them into positive, negative, or neutral sentiments (Chandra et al., 2024). This approach enables businesses to monitor customer sentiment efficiently, identify strengths and weaknesses in service delivery, and respond proactively to customer feedback (Kusumaningrum et al., 2023; Lunkes et al., 2025). In the tourism and hospitality sector, sentiment analysis has been increasingly adopted to analyze large volumes of online reviews, providing data-driven insights that would be difficult to obtain through manual analysis alone (An & Ozturk, 2022).

Several previous studies have demonstrated the effectiveness of sentiment analysis using machine learning algorithms in different contexts. Arifin & Purnama (2023) conducted sentiment analysis on Twitter data related to GrabFood and GoFood services, achieving accuracy levels above 85%. Nata & Maarif (2024) analyzed customer sentiment toward locally produced fashion products sold online and found that more than 70% of topics reflected positive sentiment. In the tourism context, Hidayat et al. (2021) applied the Support Vector Machine algorithm to analyze Twitter sentiment about Rinca Island and achieved an accuracy of 87%. Although these studies confirm the reliability of machine learning approaches, most focus on social media platforms or non-hotel services. Research that specifically analyzes online hotel reviews in localized marine tourism destinations such as Senggigi Beach, particularly through a comparative evaluation of machine learning algorithms, remains limited. This gap highlights the need for more context-specific studies in tourism analytics.

Based on this gap, this study aims to conduct sentiment analysis on online hotel reviews in the Senggigi Beach area using two popular machine learning algorithms: Decision Tree and Support Vector Machine. The objectives of this research are to identify dominant customer sentiments expressed in hotel reviews and to compare the performance of both algorithms in classifying sentiment accurately. Academically, this study contributes to the literature by extending sentiment analysis research to a specific marine tourism destination and providing a comparative evaluation of machine learning methods in hotel review analysis. Practically, the findings offer valuable insights for hotel managers and policymakers in NTB Province to improve service quality, enhance customer satisfaction, and support data-driven decision-making in tourism development.

B. METHODS

This research is a quantitative study that applies a comparative machine learning approach to analyze sentiment in online hotel reviews. The performance of Decision Tree and Support Vector Machine algorithms is evaluated and compared using classification accuracy. The research methodology used in this study is illustrated in Figure 1. The initial stage involves data collection to obtain the dataset. Next, data preprocessing is performed on the collected dataset. This is followed by feature weighting and modeling using several ML algorithms. The final stage is evaluation, which is conducted to assess the effectiveness of the labeling methods used.

1. Data Collection

The dataset used in this study consists of online hotel reviews collected from the TripAdvisor platform, focusing on hotels located in the Senggigi Beach area, Lombok, West Nusa Tenggara. Data collection was conducted using a web scraping technique implemented through Google Colab with Python-based scraping libraries. The scraping process was carried out over a defined period by retrieving reviews published between 2006–2024. The dataset includes reviews from several well-known hotels in the Senggigi area, such as Holiday Resort Hotel, Merumatta Senggigi Lombok, Qunci Villas Resort, Aruna Senggigi Resort, and Sudamala Resort. A total of 6,920 review records written in various languages were successfully collected. Each data record contains the following attributes: review ID, review title, review text, rating score, and trip year. Only English-language reviews were included to maintain consistency in text processing and analysis.

2. Preprocessing

Preprocessing is the process of preparing raw data before further processing. This step is essential for improving data quality and consistency by eliminating issues that could interfere with data processing (Chi et al., 2025; Paneru et al., 2025; Taher Karim, 2024). The preprocessing steps such as data cleaning, case folding, tokenization, stopword removal, and stemming (Fakhrezi et al., 2023; Haziq et al., 2024; Nata & Maarif, 2024). The explanation follows:

a. Duplicate Removal: Identical reviews were removed to avoid bias.
b. Data Cleaning: Special characters, punctuation, numbers, URLs, and emojis were removed.
c. Case Folding: All text was converted to lowercase.
d. Tokenization: Text was split into individual tokens (words).
e. Stopword Removal: Common English stopwords were removed using a predefined stopword list. Stemming: Words were reduced to their root form to minimize vocabulary size.
f. Labeling is the process of assigning sentiment labels to dataset. Sentiment is categorized into three labels: positive, negative, and neutral. Labeling conducting with review's rating by converting ratings of 1 or 2 into negative sentiment, a rating of 3 into neutral sentiment, and ratings of 4 or 5 into positive sentiment.

3. Vectorization

After preprocessing, the text-based dataset is converted into numerical format, a process known as vectorization. Vectorization is performed using CountVectorizer. CountVectorizer is a class in the scikit-learn library that transforms a collection of text documents into a numerical matrix based on word or token frequency (Deepthi et al., 2025; Goyal, 2021). This class includes several parameters that assist in text preprocessing, such as stopword removal, word count limits (maximum and minimum), vocabulary restrictions, n-gram generation, and more (Arifudin et al., 2023). In this study, vectorization was applied with default parameter settings, without any additional configuration or parameter tuning.

4. Modeling

This study employs two ML algorithms, namely Decision Tree, and Support Vector Machine. The Decision Tree algorithm constructs a tree-based structure for classification, while SVM identifies an optimal hyperplane to separate sentiment classes in high-dimensional space. A detailed explanation of Decision Tree is discussed in (Sinha et al., 2024),(Cam et al., 2024), and Support Vector Machine in (Cam et al., 2024; Nurhaliza Agustina et al., 2024). During this stage, the dataset is split into two parts: 80% for training data and 20% for testing data. All models were implemented using the Python programming language and the Scikit-learn library.

5. Evalaution

After the modeling stage, the next step is to evaluate the sentiment analysis results for each labeling method using accuracy. Accuracy can be calculated using Equation (1) (Ramadhan et al., 2022; P. Singh et al., 2021). In addition to accuracy, precision, recall, and F1-score were employed to provide a more comprehensive evaluation of the classification performance.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{1}$$

where:

TP (True Positive) : The condition where the ML correctly predicts the positive class.

TN (True Negative) : The condition where the ML correctly predicts the negative class.

FP (False Positive) : The condition where the ML incorrectly predicts the positive class, while it should be negative.

FN (False Negative) : The condition where the ML incorrectly predicts the negative class, while it should be positive.

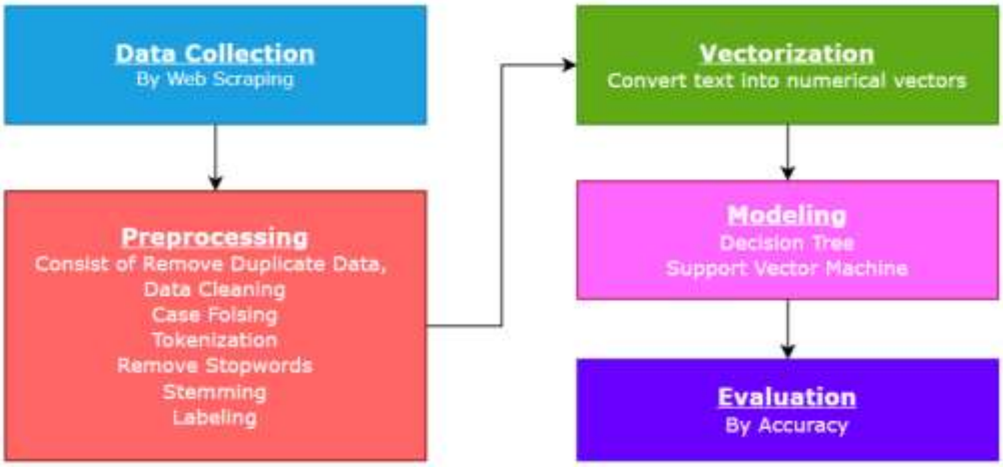

Figure 1. Methodology

C. RESULT

1. Data Collection

The dataset consist of 6,920 data were successfully collected by web scraping. An overview of the dataset is presented in Figure 2. For sentiment analysis, the columns used are "Review" and "Rate", while the "Id", "Title" and "Trip Year" columns are removed.

	Id	Title	Review	Rate	Trip Year
0	749407036	Beach Front	We took a villa with private pool. The room is...	4	February 2020
1	745822196	Absolutely magnificent!!! Go there!	We had a absolutely magnificent time in Sudama...	5	February 2020
2	741855103	Exceptional small resort	Had awesome stay at this lovely resort. Was re...	5	January 2020
3	746768285	Simply wow!	This place is amazing! Our room was lovely wit...	5	February 2020
4	751163541	Absolute Perfection	Wow, my partner and I are already planning our...	5	February 2020
...	...	...	...	...	...
6915	910439467	The best, most comfortable resort in the Mangs...	I'm really happy to be able to vacation in a p...	5	August 2023
6916	909954163	Good for chilling by the beach	Family hotel very good for chilling by the bea...	4	August 2023
6917	909888320	Days on Lombok	The hotel is beautifully situated on the beach...	4	August 2023
6918	908487431	Holiday experience	I'm so happy, this holiday with my friends, wh...	5	July 2023
6919	907269209	An Exciting Holiday that is Comfortable and Safe	On holiday with your family, especially if you...	5	June 2023

6920 rows × 5 columns

Figure 2. Dataset

2. Preprocessing

Preprocessing is performed to enhance the dataset quality (Velmala et al., 2025). During this stage, only English-language reviews were retained for analysis, resulting in a final dataset of 6,864 data points. Labeling resulted in 6,268 positive, 389 negative, and 207 neutral data. The results from this process will be converted into numerical format and used as training and testing data for machine learning modeling. An overview of this process can be seen in Figure 3.

	Review	Rate	Rating_Label
0	took villa privat pool room medium size big be...	4	Positive
1	absolut magnific time sudamala resort lombok r...	5	Positive
2	awesom stay love resort realli special staff b...	5	Positive
3	place amaz room love great view pool sea nice ...	5	Positive
4	wow partner alreadi plan next stay hotel	5	Positive
...	...	...	...
6915	im realli happi abl vacat place like especi si...	5	Positive
6916	famili hotel good chill beach swim pool	4	Positive
6917	hotel beauti situat beach room well furnish go...	4	Positive
6918	im happi holiday friend arriv hotel immedi gre...	5	Positive
6919	holiday famili especi small children realli ch...	5	Positive

6864 rows × 3 columns

Figure 3. Preprocessing Results

3. Sentiment Analysis

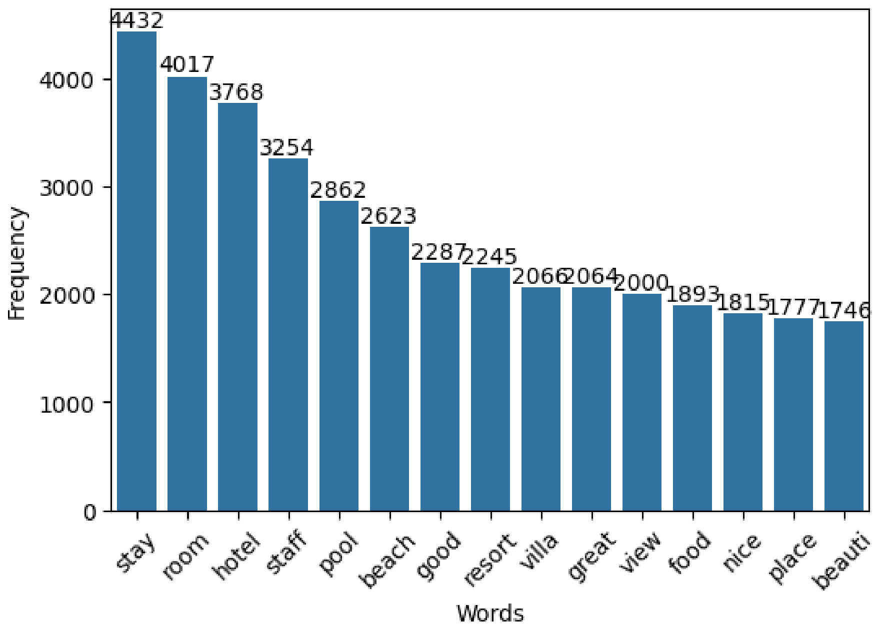

Figure 4. Positive Word Frequency

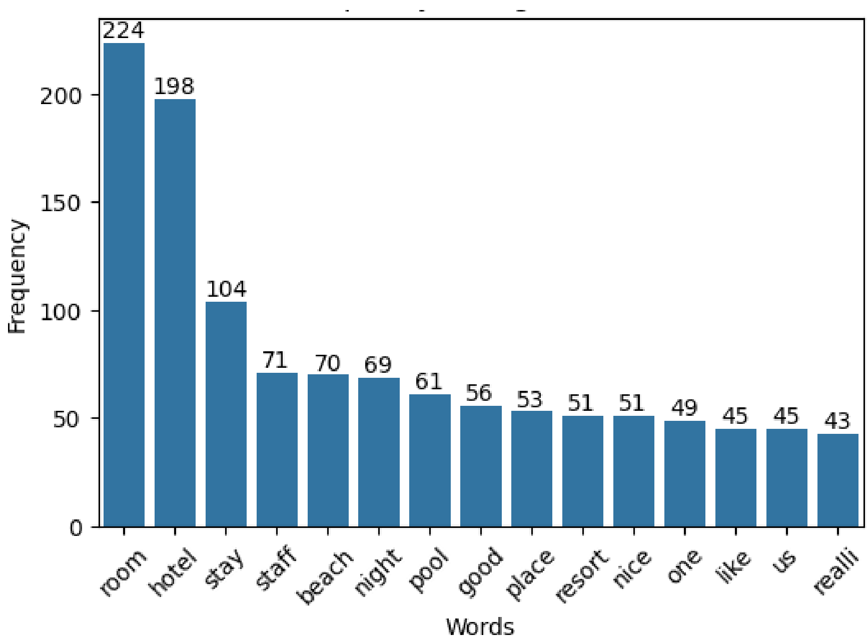

Figure 5. Negative Word Frequency

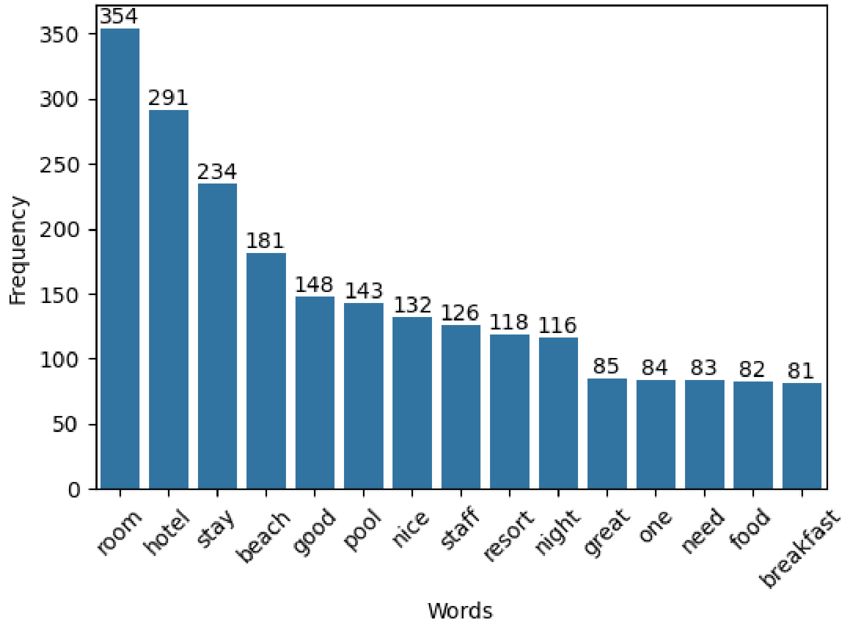

Figure 6. Neutral Word Frequency

Figure 4, Figure 5, and Figure 6 illustrates the word frequency distribution in sentiment analysis where Figure 4 for positive sentiment, Figure 5 for negative sentiment, and Figure 6 for neutral sentiment. Based on the Figures, the words “room”, “hotel”, and “stay” appear across all sentiment categories, emphasizing that accommodation experience and room quality significantly influence customer satisfaction. Positive sentiment predominantly features adjectives such as “great”, “nice’ and “beautiful” indicating a pleasant experience. In negative sentiment, the sentiment labeling process did not identify explicitly negative words, as reviews on the platform tend to use more polite expressions like “not good”, ‘not very nice”, or “don’t like” rather than directly using the word “bad”. Negative sentiment is primarily associated with staff and hotel environment, highlighting that the main factors contributing to negative reviews are facility or service-related issues. Meanwhile, neutral sentiment tends to be descriptive, focusing more on experiential aspects without expressing strong emotions, as shown in Table 1.

Table 1. Evaluation of Machine Learning Algorithms

Metrics	Decision Tree	Support Vector Machine
Accuracy	87%	90%
Precision	84%	91%
Recall	87%	90%
F1-Score	85%	85%

Table 1 presents the evaluation results of the Decision Tree and Support Vector Machine (SVM) models using multiple performance metrics, including accuracy, precision, recall, and F1-score. Overall, the SVM model outperformed the Decision Tree across most metrics, achieving an accuracy of 90%, precision of 91%, recall of 90%, and F1-Score, compared to 87%, 84%, and 87%, respectively, for the Decision Tree. These results indicate that SVM provides a more reliable classification of sentiment in hotel reviews from the Senggigi Beach area.

The stronger performance of SVM can be technically justified by its ability to handle high-dimensional feature spaces generated through text vectorization. Using Count Vectorizer with default parameters, each unique term in the corpus was represented as a feature, resulting in sparse and high-dimensional input data. SVM is well suited for this type of data, as it constructs an optimal hyperplane that maximizes the margin between sentiment classes, enabling better generalization when dealing with diverse vocabulary and complex sentence structures. In contrast, the Decision Tree model, while interpretable, is more sensitive to noise and data sparsity, which may reduce its effectiveness in text-based sentiment classification.

The results of this study are consistent with and support findings from previous research on sentiment analysis. The accuracy achieved by the Support Vector Machine (90%) is higher than that of the Decision Tree (87%), which aligns with earlier studies that reported superior performance of SVM in text-based sentiment classification. For example, (Hidayat et al., 2021) found that SVM achieved an accuracy of 87% in analyzing Twitter data related to tourism, confirming its robustness in handling high-dimensional textual features. Similarly, (Arifin & Purnama, 2023) reported accuracy levels above 85% when applying machine learning models for sentiment analysis of online service reviews, indicating that ML-based approaches are effective in capturing customer sentiment. Compared to these studies, the accuracy obtained in this research is slightly higher, suggesting that SVM remains highly reliable when applied to hotel review data in the tourism domain. Overall, the findings support previous research by reaffirming that SVM consistently outperforms simpler classification models, such as Decision Tree, while both algorithms are capable of producing strong and reliable sentiment classification results.

4. CONCLUSION

Sentiment analysis of hotel reviews helps users understand the condition of a hotel. Additionally, it enables hotel management to identify their strengths and weaknesses, allowing them to make informed decisions for service improvement. The results indicate that Positive sentiment reflects a pleasant experience. Negative sentiment indicates that the main factors in negative reviews are issues related to facilities or services. Neutral sentiment tends to be descriptive, primarily discussing experiences without strong emotional expressions. The results demonstrate that both algorithms are capable of effectively classifying customer sentiment. The Decision Tree model achieved an accuracy of 87%, precision of 84%, recall of 87%, and an F1-score of 85%. Meanwhile, the SVM model outperformed Decision Tree across most evaluation metrics, obtaining an accuracy of 90%, precision of 91%, recall of 90%, and an F1-score of 85%. These findings confirm that SVM provides more robust and consistent performance for sentiment classification in hotel review data.

From an academic perspective, this research contributes empirical evidence to the field of tourism analytics by demonstrating the comparative effectiveness of machine learning algorithms using multiple evaluation metrics rather than accuracy alone. The inclusion of precision, recall, and F1-score highlights the importance of balanced evaluation, particularly in datasets dominated by positive sentiment. This study reinforces prior research indicating that SVM consistently outperforms simpler classifiers in text-based sentiment analysis while maintaining stable classification performance. Practically, the results offer valuable insights for hotel managers and tourism policymakers in West Nusa Tenggara Province. By focusing on improving room quality, staff performance, and facility management, hotels can directly address the most frequently mentioned factors in negative reviews. Monitoring online reviews through sentiment analysis enables proactive service improvement, which can enhance tourist satisfaction. Despite its contributions, this study has several limitations. First, the dataset was dominated by positive reviews, which may have influenced model performance and limited the accuracy of minority sentiment classification. Second, the analysis focused on reviews from a single online platform and used only two machine learning algorithms. Additionally, the study relied on textual content without considering contextual factors such as reviewer demographics or temporal trends. Future research is recommended to address these limitations by incorporating data from multiple review platforms, applying techniques to handle class imbalance, and exploring advanced models such as deep learning or hybrid approaches.

ACKNOWLEDGEMENT

The authors want to express our sincere gratitude to our Professor, Dr. Heri and Dr, Wira for his invaluable guidance and support throughout this research. Their expertise and insights have contributed significantly to the success of this research. We would also like to thank the University of Mataram for providing this study's necessary resources and facilities.

REFERENCES

An, Q., & Ozturk, A. B. (2022). Assessing the Effects of User-Generated Photos on hotel Guests' Price, Service Quality, Overall Image Perceptions and Booking Intention. *Journal of Hospitality and Tourism Technology*, *13*(4), 608–625. https://doi.org/10.1108/JHTT-05-2021-0146

Arifin, R., & Purnama, D. A. (2023). Identifying customer preferences on two competitive startupproducts: An analysis of sentiment expressions and textmining from Twitter data. *Jurnal Infotel*, *15*(1), 66–74. https://doi.org/10.20895/infotel.v15i1.906

Arifudin, R., Indra Wijaya, D., Warsito, B., & Wibowo, A. (2023). Voting Classifier Technique and Count Vectorizer with N-gram to Identify Hate Speech and Abusive Tweets in Indonesian. *Scientific Journal of Informatics*, *10*(4), 469. https://doi.org/10.15294/sji.v10i4.46633

Burhanudin, B. (2024). Managing social commerce: does customer review quality matter? *Procedia Computer Science*, *234*, 1459–1466. https://doi.org/10.1016/j.procs.2024.03.146

Cam, H., Cam, A. V., Demirel, U., & Ahmed, S. (2024). Sentiment analysis of financial Twitter posts on Twitter with the machine learning classifiers. *Heliyon*, *10*(1), e23784. https://doi.org/10.1016/j.heliyon.2023.e23784

Chandra, R., Zhu, B., Fang, Q., & Shinjikashvili, E. (2024). Large language models for sentiment analysis of newspaper articles during COVID-19: The Guardian. *Applied Soft Computing*, *171*(October 2024), 112743. https://doi.org/10.1016/j.asoc.2025.112743

Chi, D., Huang, T., Jia, Z., & Zhang, S. (2025). Research on sentiment analysis of hotel review text based on BERT-TCN-BiLSTM-attention model. *Array Journal*, *25*(1), 1–10. https://doi.org/10.1016/j.array.2025.100378

D. Deepthi, Kamineni B. T. Sundari, B. Satish BABU, K.Thrilochana, A.Mahalakshmi, & Sneha.H.Dhoria. (2025). Leveraging Count Vectorizer For Job Title Prediction: A Comparative Study Of Machine Learning Algorithms. *Journal of Theoretical and Applied Information Technology*, *15*(5), 1781–1794. http://www.jatit.org/volumes/Vol103No5/12Vol103No5.pdf

Deepthi, S., & Shariff, D. (2024). Role Of Hotel Industry In Tourism Development. *Educational Administration: Theory and Practice*, *30*(5), 10088–10091. https://doi.org/10.53555/kuey.v30i5.4347

Fakhrezi, M. F., Adian Fatchur Rochim, & Dinar Mutiara Kusomo Nugraheni. (2023). Comparison of Sentiment Analysis Methods Based on Accuracy Value Case Study: Twitter Mentions of Academic Article. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, *7*(1), 161–167. https://doi.org/10.29207/resti.v7i1.4767

Goyal, R. (2021). Evaluation of rule-based, CountVectorizer, and Word2Vec machine learning models for tweet analysis to improve disaster relief. *2021 IEEE Global Humanitarian Technology Conference (GHTC)*, 16–19. https://doi.org/10.1109/GHTC53159.2021.9612486

Hasibuan, I. M., Mutthaqin, S., Erianto, R., & Harahap, I. (2023). Contribution of the Tourism Sector to the National Economy. *Jurnal Masharif Al-Syariah: Jurnal Ekonomi Dan Perbankan Syariah*, *8*(2), 1200–1217. https://doi.org/http://dx.doi.org/10.30651/jms.v8i2.19280

Haziq, M. R., Sibaroni, Y., & Prasetyowati, S. S. (2024). Word Embedding Optimization In Sentiment Analysis Of Reviews On Mytelkomsel App Using Long Short-Term Memory And Synthetic Minority Over-Sampling Technique. *Jurnal Teknik Informatika (JUTIF)*, *5*(6), 1581–1589. https://doi.org/10.52436/1.jutif.2024.5.6.2498

Hidayat, T. H. J., Ruldeviyani, Y., Aditama, A. R., Madya, G. R., Nugraha, A. W., & Adisaputra, M. W. (2021). Sentiment analysis of twitter data related to Rinca Island development using Doc2Vec and SVM and logistic regression as classifier. *Procedia Computer Science*, *197*, 660–667. https://doi.org/10.1016/j.procs.2021.12.187

Islahuddin, Muhtasom, A., Masatip, A., & Herman, H. (2024). Enhancing Sustainable Marine Tourism: The Crucial Role of Environmental Conservation Programs in West Nusa Tenggara. *Proceedings International Conference on Marine Tourism and Hospitality Studies*, *1*(1), 100–107. https://doi.org/10.33649/iconmths.v1i1.369

Jamal Ali, B., Gardi, B., Jabbar Othman, B., Ali Ahmed, S., Burhan Ismael, N., Abdalla Hamza, P., Mahmood Aziz, H., Yassin Sabir, B., Sorguli, S., & Anwar, G. (2021). Hotel Service Quality: The Impact of Service Quality on Customer Satisfaction in Hospitality. *International Journal of Engineering, Business and Management (IJEBM)*, *5*(3), 14–28. https://doi.org/10.22161/ijebm.5.3

Kusumaningrum, R., Nisa, I. Z., Jayanto, R., Nawangsari, R. P., & Wibowo, A. (2023). Deep learning-based application for multilevel sentiment analysis of Indonesian hotel reviews. *Heliyon*, *9*(6), e17147. https://doi.org/10.1016/j.heliyon.2023.e17147

Lunkes, R. J., Codesso, M., da Rosa, F. S., Mendes, A. C., & Costa, G. D. (2025). How Managers' Perceptions Of Online Reviews Enhance Digital Innovation in Hotels: The role of Technological Opportunism. *International Journal of Hospitality Management*, *133*(2), 104416. https://doi.org/10.1016/j.ijhm.2025.104416

Nata, I. A., & Maarif, M. R. (2024). Understanding Customer Perception of Local Fashion Product on Online Marketplace through Content Analysis. *Jurnal Infotel*, *16*(1), 58–70. https://doi.org/10.20895/infotel.v16i1.1070

Nurhaliza Agustina, C. A., Novita, R., Mustakim, & Rozanda, N. E. (2024). The Implementation of TF-IDF and Word2Vec on Booster Vaccine Sentiment Analysis Using Support Vector Machine Algorithm. *Procedia Computer Science*, *234*, 156–163. https://doi.org/10.1016/j.procs.2024.02.162

Paneru, B., Thapa, B., & Paneru, B. (2025). Sentiment analysis of movie reviews: A flask application using CNN with RoBERTa embeddings. *Systems and Soft Computing*, *7*(September 2024), 200192. https://doi.org/10.1016/j.sasc.2025.200192

Ramadhan, N. G., Wibowo, M., Mohd Rosely, N. F. L., & Quix, C. (2022). Opinion mining indonesian presidential election on twitter data based on decision tree method. *Jurnal Infotel*, *14*(4), 243–248. https://doi.org/10.20895/infotel.v14i4.832

Singh, P., Singh, N., Singh, K. K., & Singh, A. (2021). Chapter 5 - Diagnosing of disease using machine learning. In K. K. Singh, M. Elhoseny, A. Singh, & A. A. Elngar (Eds.), *Machine Learning and the Internet of Medical Things in Healthcare* (pp. 89–111). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-821229-5.00003-3

Sinha, A., Rout, B., Mohanty, S., Mishra, S. R., Mohapatra, H., & Dey, S. (2024). Exploring Sentiments in the Russia-Ukraine Conflict: A Comparative Analysis of KNN, Decision Tree and Logistic Regression Machine Learning Classifiers. *Procedia Computer Science*, *235*, 1068–1076. https://doi.org/10.1016/j.procs.2024.04.101

Suhono, H. A. R., Pratiwi, R. A., & Kurniadhi, A. (2020). GIS-based environmental assessment of selected prioritized tourist attractions on Lombok Island. *IOP Conference Series: Earth and Environmental Science*, *592*(1), 1–37. https://doi.org/10.1088/1755-1315/592/1/012014

Taher Karim, S. H. (2024). Kurdish social media sentiment corpus: Misyar marriage perspectives. *Data in Brief*, *57*(12), 110989. https://doi.org/10.1016/j.dib.2024.110989

Velmala, M., Rajiakodi, S., Pannerselvam, K., & Sivagnanam, B. (2025). Multimodal Sentiment Analysis of Online Memes: Integrating Text and Image Features for Enhanced Classification. *Procedia Computer Science*, *258*, 355–364. https://doi.org/10.1016/j.procs.2025.04.272

Wahyuningtyas, N., Tanjung, A., Kodir, A., & Wijanarko, H. (2020). Management of Tourism Areas Based on Disaster Mitigation (Case Study of Senggigi Beach). *IOP Conference Series: Earth and Environmental Science*, *412*(1), 1–6. https://doi.org/10.1088/1755-1315/412/1/012015