

Strategic Action for Enhancing Bali Tourism Based on Sentiment Analysis and Topic Modeling

Muhamad Syukron
Department of Statistics
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia
Muhamad.syukron@its.ac.id

Jahidata Royyanida Mardani
Department of Statistics
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia
5003231032@student.its.ac.id

Adelia Tria Amanda
Department of Statistics
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia
5003231038@student.its.ac.id

Tesalonika Sharavova Hutagaol
Department of Statistics
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia
5003231033@student.its.ac.id

Article History

Received April 9th, 2026

Revised May 16th, 2026

Accepted July 15th, 2026

Published August, 2026

Abstract— Bali was a prominent global tourism destination and a leading contributor to Indonesia's tourism revenue. Despite its unique attractions, Bali faced challenges such as overtourism, traffic congestion, and environmental degradation. To explore ways to improve Bali's tourism, this study analyzed visitor suggestions derived from extensive Google Maps reviews, which were too numerous to be processed manually. A deep learning approach was employed by integrating textual reviews and user-attached images through a multimodal analysis framework. The BLIP Visual Language Model was used to generate captions from tourism images, which were then concatenated with the corresponding textual reviews. The combined text was processed using Sentence BERT embeddings and classified into positive, neutral, and negative sentiments. Subsequently, BERTopic was applied to identify the main discussion themes and generate actionable recommendations. The results revealed nine key topics, including Environmental Quality and Comfort of Tourist Sites, Tourism Infrastructure and Accessibility, and Tourism Safety. The findings provided strategic insights for tourism authorities and stakeholders to identify priority issues, support tourism improvement planning, allocate development resources more effectively, and enhance visitor experience in Bali. In addition, this study demonstrated how multimodal sentiment analysis supported data-driven decision-making for tourism management and policy development.

Keywords—BERTopic; BLIP; multimodal analysis; sentence BERT embedding; textual reviews

1 INTRODUCTION

Bali's reputation as one of the world's premier tourism destinations is firmly rooted in its rich cultural heritage, stunning natural landscapes, and vibrant local traditions [1], [2], [3]. The island's unique combination of pristine beaches, terraced rice fields, and spiritual sites has attracted millions of visitors annually, solidifying its pivotal role in Indonesia's tourism industry and the global travel market [4], [5], [6]. Extensive tourism development, including accommodations and infrastructure, has propelled Bali's status as a major tourist hotspot [7], [8]. Efforts to balance economic benefits with environmental preservation and respect for local culture underscore Bali's ongoing challenges and opportunities in tourism development [9], [10], [11], [12].

Despite its continued development, Bali has recently faced several tourism-related challenges, including overcrowding, environmental degradation, and problematic tourist behaviors [5], [13], [14]. Public perceptions and discourse analysis of visitor experiences highlight mixed sentiments where tourists often appreciate Bali's cultural landmarks and natural beauty but their activities contribute to overcrowding and environmental strain, particularly in hotspot areas like Ubud and Canggu [15], [16], [17]. Problematic tourist behaviors, including littering and disrespect for local customs, further complicate conservation efforts [9].

Managing the challenges in sustainable tourism necessitates targeted and data-driven strategies [18], [19], [20], [21]. Advances in technology have made vast amounts of data available, such as Google Maps reviews of tourist sites, which can be leveraged effectively through sentiment analysis and topic modeling. Research demonstrates that sentiment analysis is a potent tool in the tourism industry for extracting subjective data concerning traveler satisfaction, preferences, and perceptions [22], [23]. Likewise, topic modeling methods such as Latent Dirichlet Allocation (LDA) help classify user reviews by uncovering latent topics and patterns relating to aspects like experience, transportation, natural scenery, and service management, as shown in-depth analyses of tourist destinations [24], [25].

The utilization of sentiment analysis and topic modeling has been extensively documented within tourism research as effective methods for extracting insights from textual data such as online reviews [26], [27]. However, the potential of leveraging image data, particularly user-uploaded images on platforms like Google Maps [28], despite its capacity to offer complementary and richer insights into tourist experiences. Therefore, this research develops a multimodal sentiment analysis and topic modeling framework to identify the main suggestions and concerns expressed by visitors regarding tourism in Bali. The proposed framework is intended to support tourism stakeholders and policy-makers in identifying priority issues, understanding visitor perceptions, planning tourism improvement strategies, and allocating development resources more effectively based on data-driven insights. In addition, this study contributes to tourism analytics research by demonstrating the integration of textual

and visual information within a multimodal sentiment analysis framework for tourism decision support.

2 METHOD

2.1 Literature Review

Sentiment analysis has emerged as a crucial tool in understanding tourists' perceptions, emotions, and satisfaction through their online reviews and social media posts. The study by [24] conducted sentiment analysis using SnowNLP on nearly 27,000 reviews of tourist attractions in Guizhou Province, China. It revealed a generally positive sentiment trend among tourists and identified six key attributes influencing satisfaction, including experience and transportation, particularly highlighting their importance in ethnic tourism. This work provides empirical insights that supplement traditional surveys by capturing diverse and dynamic tourist opinions.

Similarly, [29] applied sentiment analysis to Twitter data related to metaverse applications in the hospitality sector. Their findings showed a predominance of positive sentiments and uncovered relevant topics such as gaming and virtual sightseeing, illustrating the capacity of sentiment analysis and topic modeling to capture evolving customer experiences and identify emerging concerns in tourism-related online content. Although the study focused on metaverse applications rather than conventional tourism destinations, it demonstrates the applicability of these text-mining techniques to extracting insights from user-generated content in the tourism and hospitality domain.

Topic modeling complements sentiment analysis by uncovering latent themes and attributes from unstructured textual data, thus enriching the understanding of tourist experiences and preferences. In [24], Latent Dirichlet Allocation (LDA) was applied to extract topics related to six attributes influencing tourist satisfaction. This combination of sentiment and topic analyses offers a holistic perspective, revealing temporal trends and distinctions across tourist groups and advancing the ethnic tourism literature.

Ref. [29] utilized topic modeling to identify six important topics such as virtual events and business within metaverse applications in hospitality, demonstrating how topic models help detect emerging trends and customer concerns in novel tourism contexts.

2.2 Image Captioning

Image captioning is a technique that generates descriptive textual content for images, enabling the extraction of meaningful information from visual data. In this study, we employ the BlipProcessor [30], a state-of-the-art model designed for efficient and accurate image captioning tasks. BlipProcessor integrates advanced vision and language models to process images and produce coherent, contextually relevant captions.

The BlipProcessor operates by encoding the input image into a feature representation using a vision transformer,



followed by a language generation module that decodes these features into natural language descriptions. This approach allows for capturing intricate visual details and translating them into semantically rich captions, which are essential for analyzing visual content within tourism-related datasets. The architecture of this model is illustrated in Fig. 1.

By utilizing BlipProcessor, we can systematically generate captions for images associated with tourist experiences, enabling the incorporation of visual data into the overall analysis framework. This enhances the understanding of tourist preferences and behaviors by complementing textual data with detailed image descriptions.

2.3 Text Embedding

To effectively integrate visual and textual data, this study encodes both the image captions generated by BlipProcessor and the original textual reviews using a unified format. Before concatenation, all textual reviews and generated image captions underwent preprocessing, including lowercasing, removal of special characters, and whitespace normalization. Each data instance is structured by enclosing the image caption and the review text within distinct tags as follows:

```
<begin image caption> [image caption text] </end image caption> <begin review text> [review text] </end review text>
```

The structured input is subsequently processed using BERT to derive sentence embeddings. Initially, the combined tagged text is tokenized into subword units and converted into token IDs via BERT's tokenizer. Each token ID is transformed into a dense vector through an embedding layer that integrates token embeddings, positional embeddings, and segment embeddings. This integration enables the model to represent both the semantic content of words and their positional context within the sequence.

The resulting sequence of vectors is transmitted through multiple Transformer encoder layers. Within each layer, self-

attention mechanisms compute the relationships among all tokens, thereby allowing the model to capture contextual dependencies across the image captions and review texts. This process facilitates BERT's comprehension of the interrelations among different components of the input within a unified representation.

Upon traversing the Transformer layers, the final hidden states for each token are generated. To obtain a fixed-length sentence embedding, the representation of the special [CLS] token is typically employed. This vector functions as a summary of the entire input, encapsulating the integrated semantic meaning of both the visual descriptions and the user-generated textual content.

These embeddings serve as the basis for downstream tasks such as clustering, sentiment analysis, or topic modeling, facilitating a more comprehensive understanding of tourist experiences by leveraging multimodal information. The text embedding process is shown in Fig. 2.

2.4 Sentiment Classification Model

The encoded embeddings from the combined image captions and textual reviews, each of dimension 768, serve as input to a neural network classifier designed for multi-class classification. The architecture consists of an input layer of size 768, followed by a hidden layer with 256 neurons, and an output layer with 3 neurons corresponding to the target classes.

The model was trained from scratch, with its parameters randomly initialized and optimized during training without using pretrained weights, using the cross-entropy loss function, which is well-suited for classification tasks with mutually exclusive classes. This setup enabled the network to learn discriminative features from the integrated multimodal embeddings, facilitating the categorization of tourist experiences, as formulated in (1). Fig. 3 presents the architecture of the neural network classifier.

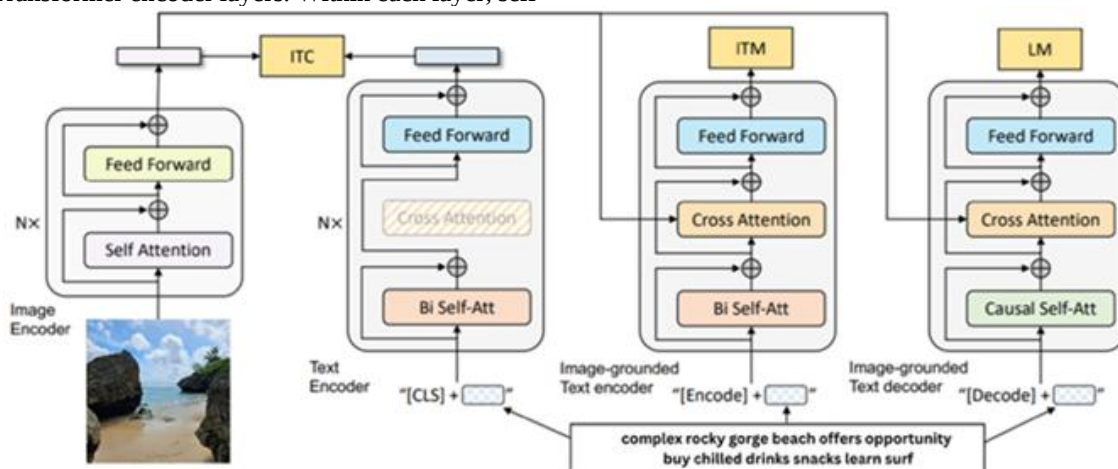


Figure 1. BLIP architecture



$$\mathcal{L} = - \sum_{c=1}^C y_c \log(p_c)$$

where:

$y_c = 1$ if class c is correct, 0 otherwise (one-hot encoding)

p_c = predicted probability for class c

[0.933, 0.111, 0.315, 0.226, 0.822, 0.767, 0.112, 0.331, 0.879, ..., 0.876]

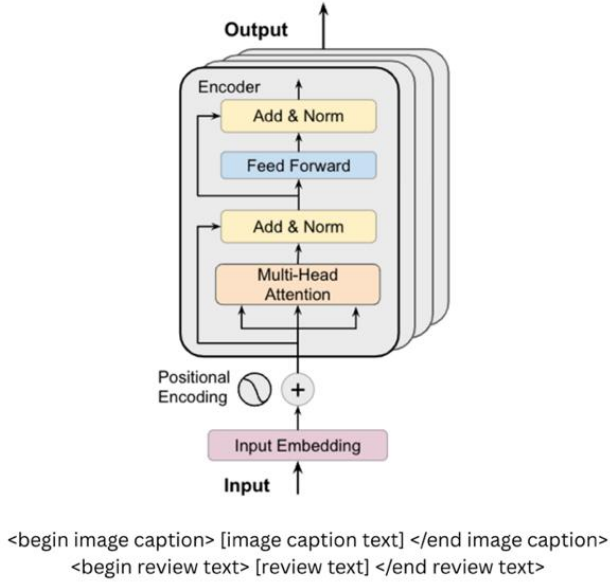


Figure 2. Text embedding

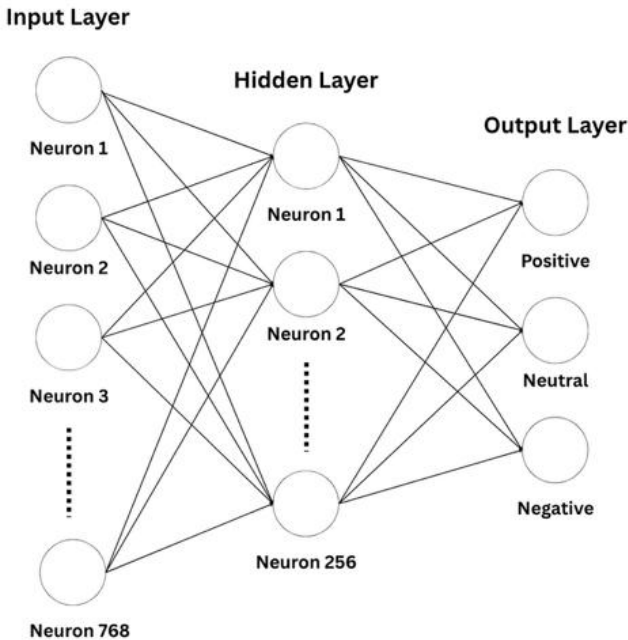


Figure 3. Neural network classifier architecture

2.5 BERTopic

In addition to sentiment analysis, this study employs BERTopic for topic modeling to uncover latent themes within the textual data. BERTopic [31] is a state-of-the-art topic modeling technique that leverages transformer-based embeddings combined with clustering algorithms to identify coherent and interpretable topics from large text corpora.

The process begins by encoding the textual reviews and image captions (already embedded via BERT as described in Section 3.2) into high-dimensional vector representations. These embeddings are then clustered using algorithms such as HDBSCAN to group semantically similar documents. Finally, a class-based TF-IDF procedure extracts representative keywords for each cluster, enabling clear topic interpretation.

By integrating BERTopic, the study captures nuanced thematic structures in tourist-generated content, providing insights beyond sentiment polarity. This approach facilitates the detection of emerging trends, key concerns, and attribute-specific discussions within the tourism domain, complementing the sentiment and classification analyses.

2.6 Silhouette Score

To evaluate the quality and coherence of the topics generated by BERTopic, this study employs the Silhouette Score as a key metric. The Silhouette Score measures how well each data point fits within its assigned cluster compared to other clusters, indicating cluster separation and compactness, and is calculated using (2).

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (2)$$

where:

$a(i)$ = minimum average distance from point i to all points in the nearest different cluster

$b(i)$ = minimum average distance from point i to all points in the nearest different cluster

Calculated for each document embedding, the score ranges from -1 to 1, where higher values signify that the document is appropriately clustered with similar items and distinctly separated from other clusters. A positive Silhouette Score thus reflects well-defined and meaningful topic groupings, enhancing confidence in the interpretability and reliability of the topic modeling results.

By incorporating the Silhouette Score in the evaluation framework, the study ensures that the identified topics are not only semantically coherent but also statistically robust, supporting more insightful analyses of tourist-generated content.



3 RESULTS AND DISCUSSION

3.1 Dataset

Tourism review data and associated images were collected from Google Maps reviews related to 11 popular tourist destinations in Bali: Nusa Dua Beach, Ulun Danu Beratan Temple, Uluwatu Temple, Garuda Wisnu Kencana (GWK), Kuta Beach, Padang Padang Beach, Seminyak Beach, The Amazing Taman Safari Bali, Penglipuran Village, Sacred Monkey Forest Sanctuary, and Melasti Beach. The data collection process was conducted using a web scraping pipeline implemented in Python with the Selenium framework. Reviews and user-uploaded images were collected within the date range from 1 April 2024 to 30 April 2026. This study did not perform sampling, as all accessible reviews containing image attachments within the specified period were collected and included in the dataset. In total, 5,713 reviews with image attachments were obtained.

For each review, image captioning was performed using the BLIP model to generate textual descriptions from user-uploaded tourism images. The generated captions were manually inspected to ensure semantic consistency with the uploaded images and their tourism-related contexts. Although this study did not conduct an independent quantitative evaluation, the reliability of the selected BLIP model is supported by its evaluation in the original BLIP study, which demonstrated competitive or state-of-the-art performance across several vision-language benchmarks compared with existing models, including CLIP, ALIGN, ALBEF, and UNIMO. Therefore, the generated captions were considered sufficiently reliable to support the downstream multimodal sentiment analysis process in this study.

The generated captions were concatenated with the original review text, and the combined input was encoded using a BERT-based embedding model. User ratings were converted into sentiment labels: ratings of 1–2 were assigned as negative, rating 3 as neutral, and ratings of 4–5 as positive. The dataset was split into training, validation, and test sets with proportions of 70%, 15%, and 15%, respectively. The training set was used to optimize the model, the validation set was used for early stopping and model selection, and the test set was used for final evaluation on unseen data.

3.2 Sentiment Analysis

Table 1 illustrates the sentiment analysis model's performance across the training, validation, and test datasets. On the training data, the model demonstrates flawless performance, achieving 100% in precision, recall, F1-score, and accuracy, which suggests it perfectly fits the training distribution. However, on the validation set, the model's performance drops to 70% precision and around 71% for recall, F1-score, and accuracy. A similar pattern is evident on the test set, where the model attains 73% precision and 71% for recall, F1-score, and accuracy. The comparable results between the validation and test sets suggest consistent generalization on new data.

To address overfitting, a dropout rate of 30% was implemented during the training phase. Although there was consideration to increase the dropout rate, it was found that doing so adversely affected convergence and resulted in underfitting. Consequently, the current setup offers a well-balanced and effective regularization for this model and dataset. Despite a performance discrepancy between the training and evaluation sets, the alignment of validation and test outcomes indicates that the model has attained stable and reasonable generalization performance.

3.3 Topic Modelling

For topic modeling, we applied BERTopic, and hyperparameter tuning was performed to determine the optimal number of topics based on semantic distance. The silhouette score was used as the evaluation metric, where a higher score indicates better-defined and more cohesive topic clustering. The results, as shown in Fig. 4, indicate that the optimal number of topics is 22.

We applied BERTopic for topic modeling and tuned the hyperparameters to determine the optimal number of topics based on semantic distance. We evaluated the clustering quality using the silhouette score, where a higher value indicates better-defined and cohesive groups. The initial analysis identified 22 optimal topics, as shown in Fig. 4. The UMAP visualization in Fig. 5 demonstrates that most topics are well separated within the embedding space. However, we found semantic overlaps in certain areas, such as Topic 1 and Topic 2, which both address waste and cleanliness issues in Bali. We merged these overlapping topics to improve interpretability. This process finalized the model into 9 optimal topics, summarized in Table 2. The following section discusses these core themes and provides strategic suggestions to enhance the quality of tourism in Bali.

Table 1. Evaluation Metrics of Sentiment Analysis

Data	Metrics			
	Precision	Recall	F1-Score	Accuracy
Train	100%	100%	100%	100%
Val	70%	71%	71%	71%
Test	73%	71%	71%	71%

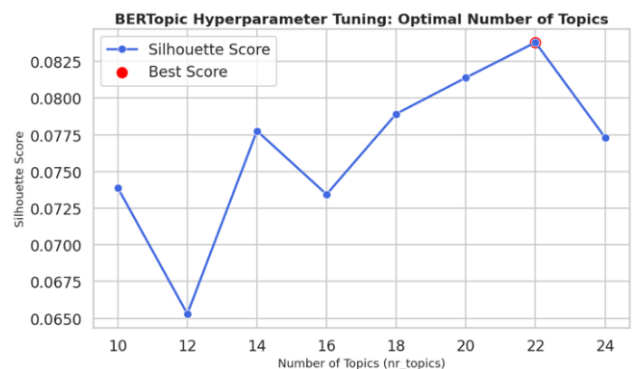


Figure 4. Silhouette score for choosing the best number of topics



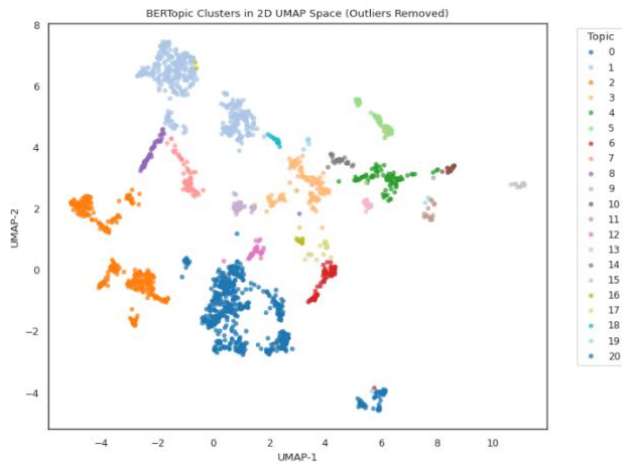


Figure 5. Cluster visualization

Table 2. Final Optimum Topics

Topic ID	Topic Name
1	Wildlife Interaction and Visitor Safety Conditions
2	Cultural-Natural Heritage and Scenic Tourism Experience
3	Traditional Village Tourism Experience
4	Environmental Quality and Comfort of Tourist Sites
5	Scenic Sunset and Time-Based Aesthetic Experience
6	Cultural Performance Tourism (Kecak Dance)
7	Tourism Infrastructure and Accessibility
8	Family-Friendly and Educational Tourism
9	Coastal Recreation and Surf Tourism

3.4 Topic 1: Wildlife Interaction and Visitor Safety Conditions

This topic addresses visitor experiences related to interactions with wildlife, specifically monkeys, and the safety concerns that arise during these encounters. It covers challenges such as protecting food and personal belongings from theft or damage by animals, as well as environmental factors like adverse weather conditions that can impact visitor safety and the overall quality of the visit. The discussion highlights the delicate balance between allowing wildlife presence and ensuring visitor well-being, as exemplified by the visitor reviews in Table 3. A positive review describes a visit where wildlife interactions are effectively managed, enabling visitors to observe monkeys safely without feeling threatened. Such experiences include clear informational signage educating visitors about animal behavior, staff monitoring the area to prevent unsafe encounters, and facilities that provide secure storage for food and valuables. In contrast, a negative review highlights situations where inadequate safety measures lead to problems such as monkeys stealing food or personal belongings, lack of supervision resulting in aggressive animal behavior, unclear or missing informational signs, and environmental challenges like unaddressed adverse weather conditions. To improve this aspect, it is recommended to implement comprehensive visitor safety management strategies. These include

establishing and enforcing clear guidelines for safe wildlife interaction, increasing the presence of trained staff in high-risk zones, and enhancing informational signage to educate visitors about animal behavior and necessary precautions. Providing secure storage options for food and belongings and developing visitor education programs focused on respectful and safe wildlife encounters will collectively enhance visitor safety and satisfaction.

3.5 Topic 2: Cultural and Natural Heritage and Scenic Tourism Experience

This topic reflects Bali's integrated cultural and natural attractions, encompassing temples, lakes, mountains, beaches, and scenic landscapes that collectively define the region's core tourism identity. It emphasizes the unique blend of cultural heritage and natural beauty that attracts visitors, highlighting the importance of preserving these assets to maintain the destination's authenticity and appeal, as illustrated by the sample reviews in Table 4. A positive review describes a visit where the cultural and natural sites are well-preserved and accessible, allowing tourists to fully appreciate the scenic beauty and heritage significance. Visitors highlight the presence of informative interpretive materials, engaging guided tours, and digital storytelling tools that enrich their understanding and connection to the sites. In contrast, a negative review points out issues such as inadequate site maintenance, lack of clear information, overcrowding, or limited access, which detract from the overall experience and diminish the appreciation of Bali's cultural and natural heritage. To improve this aspect, it is recommended to enhance site preservation efforts and maintain landscapes to prevent degradation. Providing comprehensive and engaging interpretive information through improved signage, guided tours, and digital storytelling can deepen visitor engagement and education. Developing sustainable tourism practices that balance visitor access with conservation will collectively strengthen the cultural and natural heritage experience for tourists.

3.6 Topic 3: Traditional Village Tourism Experience

This topic focuses on visitor experiences in traditional Balinese villages, highlighting key aspects such as cultural authenticity, the preservation of traditional housing structures, and maintaining cleanliness. It underscores the importance of sustaining the unique cultural identity of these villages while accommodating tourism activities. The discussion emphasizes the need to balance heritage preservation with visitor engagement to ensure that the cultural environment remains genuine and respectful, as highlighted by the visitor reviews in Table 5. A positive review describes a visit where the traditional village atmosphere is well-preserved, with authentic cultural practices and architecture clearly maintained. Visitors appreciate the cleanliness of the area and the opportunity to experience genuine local customs. Conversely, a negative review highlights issues such as commercialization that dilutes cultural authenticity, poor maintenance of traditional



structures, and inadequate sanitation. To improve this aspect, it is recommended to strengthen cultural preservation programs that protect traditional architecture and customs while promoting sustainable tourism development. Encouraging community-based tourism initiatives can enhance visitor engagement by providing authentic experiences and supporting local livelihoods. Implementing measures to maintain cleanliness and infrastructure will help ensure that the cultural environment remains attractive and respectful to both residents and visitors.



A well-maintained place in Bali, supported by nearby hotels and good facilities. The area is clean, organized, and comfortable for visitors.

Table 3. Example of Reviews from Topic 1

Example Image	Review Text
	Nice view, especially when watching the monkeys. They are cute, but they might steal glasses or phones if you are not careful.
	Went today. Entrance fee in rupiah. The monkeys are trained to steal items, especially food. One monkey even snatched my husband's sunglasses smoothly and returned them after getting food. We really underestimated their skills. The scenery is beautiful, but the monkey stealing is unforgettable.

Table 4. Example of Reviews from Topic 2

Example Image	Review Text
	Bali's beaches are beautiful and spacious, with excellent water quality. There is some sargassum seaweed. You can eat at local eateries (Warung), rent sun loungers, and take surf lessons from Balinese instructors. The beach has a friendly, relaxed atmosphere.

Table 5. Example of Reviews from Topic 3

Example Image	Review Text
	A traditional village with professional management. The houses are neatly arranged, with beautiful gapura gates and cultural ornaments. You can rent Balinese attire for photos and enjoy a beautiful temple.
	Not sure if this traditional village was created as a tourist attraction. There is an entrance fee to walk around the village, with souvenir shops and houses along the way. You mostly just walk through and are encouraged to buy things. The streets and temple area are pretty, but the many tacky shops made it less enjoyable.

3.7 Topic 4: Environmental Quality and Comfort of Tourists

This topic focuses on the environmental quality and comfort experienced by tourists at various destinations. It emphasizes key factors such as cleanliness, air quality, and the overall physical condition of the sites, which play a crucial role in shaping visitor satisfaction and well-being. The discussion highlights how maintaining a pleasant and hygienic environment is essential for enhancing the overall tourist experience and sustaining the destination's positive reputation, as reflected in the visitor reviews in Table 6. A positive review describes a visit where the tourist destination is well-maintained, featuring clean surroundings, good air quality, and comfortable facilities that contribute to a relaxing and enjoyable experience. In contrast, a negative review points out issues such as litter accumulation, unpleasant odors, poor air quality, or deteriorating infrastructure that detract from the visit. To improve this aspect, it is recommended to strengthen waste management systems through frequent collection and proper disposal methods, alongside regular site maintenance to address cleanliness and facility upkeep. Implementing environmental monitoring programs can help track air quality and other comfort

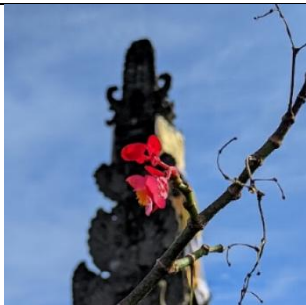


indicators, allowing timely interventions. Increasing visitor awareness campaigns to encourage responsible behaviors will support environmental preservation and cleanliness, collectively enhancing the quality and comfort of tourist destinations.

3.8 Topic 5: Scenic Sunset and Time-Based Aesthetic Experience

This topic highlights the aesthetic appreciation of sunset views and time-dependent scenic experiences, particularly in beach environments. It focuses on how visitors engage with the natural beauty of sunsets, emphasizing the importance of timing and location to fully enjoy these visual spectacles. The discussion underscores the role of environmental conditions and visitor management in shaping the overall quality and safety of the sunset viewing experience, as illustrated by the reviews in Table 7. A positive review describes a visit where sunset views are unobstructed and accessible, allowing visitors to enjoy the changing colors and ambiance in a relaxed and safe setting. Such experiences often include well-designed designated viewing platforms that provide comfortable vantage points, effective crowd management to prevent overcrowding, and amenities that enhance visitor comfort. In contrast, a negative review highlights issues such as overcrowding during peak sunset hours, lack of designated viewing areas leading to obstructed views, and safety concerns arising from unmanaged crowds or inadequate lighting. To improve this aspect, it is recommended to develop designated viewing platforms strategically located to maximize scenic appeal while ensuring visitor safety. Implementing crowd management strategies during peak sunset times can reduce congestion and enhance comfort. Providing adequate amenities and maintaining the viewing areas will support a positive visitor experience.

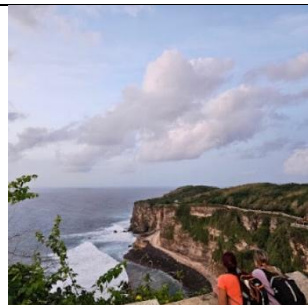
Table 6. Example of Reviews from Topic 4

Example Image	Review Text
	An honest review: the location is strategic and quite popular in Bali. The views are beautiful, and the area is clean. It is perfect for taking photos and enjoying the scenery, especially at Ulun Danu. Toilets are available if needed. Overall, it is a nice place



The cleanliness is excellent, and the buildings are very aesthetic. We visited the residents' yards, and there were many people selling knickknacks.

Table 7. Example of Reviews from Topic 5

Example Image	Review Text
	A must-visit spot. You would be sad to miss it, and we were glad we came. The sunset is beautiful, and you can watch bold monkeys near the temple. It gets dark quickly near closing time, so be careful.
	A great spot to watch the sunset. It can be a bit crowded, but it is definitely worth it for the beautiful views and atmosphere.

3.9 Topic 6: Cultural Performance Tourism (Kecak Dance)

This topic focuses on cultural performance tourism, specifically highlighting the Kecak dance at Uluwatu Temple. It explores how this traditional Balinese dance serves as both a form of cultural expression and a tourist attraction, blending entertainment with the preservation and presentation of local heritage. The discussion emphasizes the importance of managing the visitor experience to enhance cultural understanding and enjoyment, as demonstrated by the visitor reviews in Table 8. A positive review describes a visit where the Kecak dance performance is well-organized, with efficient ticketing systems that minimize wait times and comfortable seating arrangements that provide clear views of the stage. In contrast, a negative review highlights issues such as confusing or inefficient ticketing processes, overcrowded or poorly arranged seating that obstructs views, and irregular performance schedules that inconvenience visitors. To improve this aspect, it is recommended to streamline ticketing systems to ensure smooth and transparent access to performances. Optimizing seating arrangements for comfort and visibility will enhance audience engagement. Scheduling performances at regular, convenient intervals can improve



Table 9. Example of Reviews from Topic 7

Example Image	Review Text
	Padang Padang Beach is convenient, with an entrance fee in rupiah and same-day reentry. The water is clear and shallow, great for families and beginners, though there is some debris. The beach is wide, clean, and easy to access, making it a nice spot to relax.
	One of Bali's newest white sand beach destinations in the south, Melasti Beach in Ungasan is currently trending among Indonesian tourists. With road and infrastructure construction now complete, it has become a comfortable place to visit.

visitor planning and attendance. Enriching the storytelling component with detailed, multilingual explanations will deepen cultural understanding and make the experience more inclusive for diverse audiences.

3.10 Topic 7: Tourism Infrastructure and Accessibility

This topic addresses practical aspects of tourism infrastructure and accessibility, including factors such as entrance fees, parking availability, transportation access, and the overall quality of facilities. These elements are crucial for ensuring a smooth and convenient visitor experience, impacting how easily tourists can access and navigate the destination. The discussion highlights the importance of well-maintained infrastructure and efficient transportation systems to support tourism growth and visitor satisfaction, as shown in the visitor reviews in Table 9. A positive review describes a visit where tourism infrastructure is well-developed and user-friendly. Visitors experience clear and adequate signage that guides them effectively, ample and organized parking facilities, and convenient transportation options that facilitate easy access to key sites. In contrast, a negative review points out issues such as inadequate or poorly maintained facilities, limited parking availability, confusing or insufficient signage, and transportation bottlenecks that hinder smooth access. To improve this aspect, it is recommended to upgrade infrastructure quality by maintaining and expanding parking areas, improving signage for clearer navigation, and optimizing transportation flow to reduce congestion. Implementing digital ticketing systems can streamline entrance procedures, minimize waiting times, and enhance visitor convenience.

Table 8. Example of Reviews from Topic 6

Example Image	Review Text
	Loved the Kecak dance performance at Uluwatu. The artists were awesome. The tiered theater lets you watch comfortably without being disturbed by the audience. Just be careful of the monkeys.
	The venue is beautiful. Tourists should be careful with personal belongings like hats, sunglasses, food, and bags, as monkeys roam around. The Kecak dance venue is excellent, especially if you book the evening session past sunset.

3.11 Topic 8: Family-Friendly and Educational Tourism

This topic focuses on family-friendly and educational tourism, emphasizing the suitability of destinations and activities for families and children. It highlights aspects such as the educational value offered through interactive programs, the safety of the environment for younger visitors, and the availability of recreational activities that engage children and support family bonding. The discussion underscores the importance of creating tourism experiences that cater to families' needs while fostering learning and enjoyment, as illustrated by the visitor reviews in Table 10. A positive review describes a visit where the destination provides well-designed interactive educational programs that captivate children's interest and promote learning. In contrast, a negative review highlights shortcomings such as inadequate safety measures, limited or poorly maintained child-friendly facilities, and a lack of engaging educational content. To improve this aspect, it is recommended to develop more interactive and age-appropriate educational programs that actively involve children in learning experiences. Enhancing safety protocols and upgrading child-friendly infrastructure, including accessible facilities and play areas, will create a more welcoming environment. Expanding the variety of recreational activities to suit different age ranges and abilities can further support family-oriented tourism.

3.12 Topic 9: Coastal Recreation and Surf Tourism

This topic focuses on coastal recreation and surf tourism, particularly activities centered around beach environments such as surfing, wave conditions, tides, and general coastal leisure experiences. It highlights the significance of these recreational opportunities in areas like Uluwatu, emphasizing their appeal to both local and visiting enthusiasts. The



discussion also touches on the importance of balancing recreational use with safety and environmental preservation to maintain the attractiveness and sustainability of coastal tourism, as highlighted by the visitor reviews in Table 11. A positive review describes a visit where beach and surf conditions are well-managed, providing clear guidance for surfers and other visitors. In contrast, a negative review highlights issues such as insufficient safety regulations for water activities, unclear or poorly marked surfing zones that increase the risk of accidents, and inadequate beach maintenance leading to litter or environmental degradation. To enhance coastal tourism, authorities must enforce strict safety guidelines for water activities and increase lifeguard visibility to boost visitor confidence. Consistent beach maintenance and active environmental protection measures are crucial to preserve the ecosystem and ensure a safe, sustainable recreation experience.

Table 10. Example of Reviews from Topic 8

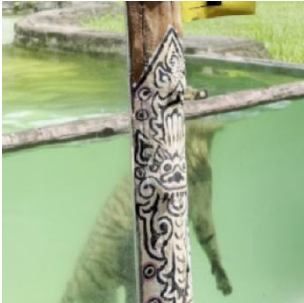
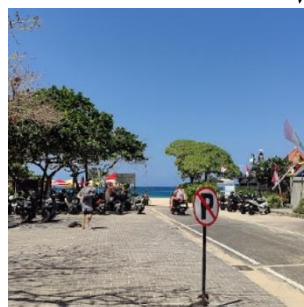
Example Image	Review Text
	Recommended kids' family friendly swim beach clean beautiful
	Bali Safari is a well-maintained, highly recommended family destination featuring diverse wildlife, scheduled shows, and a spacious food court, but photo printing is pricey.

Table 11. Example of Reviews from Topic 9

Example Image	Review Text
	This looks like a really great spot with good, gentle waves that are absolutely perfect for beginner surfers to practice on. Plus, the sunset here is incredibly pretty and worth staying for!



Overall, it is a nice beach with parking, bike rentals, and good surfing. We got an umbrella and sunbeds for 150k (down from 200k), and beach bar items average 150k. High tide trash cost it a star, but the swimming is good. For a cleaner beach with shady trees, just walk 100 meters toward the hotels and Waterblow.

4 CONCLUSION

This study proposed a multimodal framework for sentiment analysis and topic modeling of tourist reviews from major destinations in Bali. By integrating BLIP-based image captioning with textual data and encoding the combined information using BERT, the model effectively captured both visual and textual cues from user-generated content. The experimental results showed consistent performance across validation and test sets, indicating good generalization ability. Furthermore, BERTopic-based topic modeling identified 22 initial topics, which were refined into 9 meaningful clusters through topic merging based on semantic similarity and UMAP visualization. The final topics revealed key dimensions of Bali tourism, including cultural heritage, natural attractions, coastal recreation, visitor infrastructure, environmental quality, and emotional experiences. These findings demonstrated that combining multimodal sentiment analysis with topic modeling provided valuable and actionable insights for improving tourism services and enhancing visitor experience in Bali. In addition, the proposed framework could support tourism stakeholders and policymakers in identifying visitor concerns, prioritizing tourism improvement programs, supporting resource and budget allocation, and guiding data-driven tourism policies and strategic planning for sustainable tourism development in Bali. Future work may involve collaboration with the Bali tourism office and related stakeholders to validate and prioritize the proposed strategic recommendations based on resource availability, tourism development priorities, and policy constraints. Further evaluation involving domain experts may also help improve the practical applicability and decision support capability of the proposed framework.

CREDIT AUTHOR STATEMENT

Muhamad Syukron – Conceptualization, methodology design, data analyst, supervision of research activities, writing – original draft preparation, and final manuscript editing. Jahidata Royyanida Mardani – data collection, initial analysis, original draft preparation. Adelia Tri Amanda – data collection, initial analysis, original draft preparation. Tesalonika Sharavova Hutagaol – data collection, initial analysis, original draft preparation.



COMPETING INTERESTS

The authors declare that there are no competing interests or conflicts of interest related to the publication of this paper.

DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

The authors declare that generative AI and AI-assisted technologies were used in the preparation of this manuscript solely for improving language clarity, grammar, and formatting. All intellectual content, analysis, interpretations, and conclusions are the original work and responsibility of the authors.

REFERENCES

- [1] N. N. S. Astuti, P. T. Virginiya, N. K. Bagiastuti, and E. Septevany, "The future of green tourism in Bali: Challenges and opportunities," *Journal of Applied Sciences in Travel and Hospitality*, vol. 7, no. 2, pp. 178–195, Sep. 2024, doi: 10.31940/jasth.v7i2.178-195.
- [2] Kadek Wiweka and Sylvine Pickel-Chevalier, "Bali Tourism Research Trends: A Systematic Review, 1976–2022," *Jurnal Kajian Bali (Journal of Bali Studies)*, vol. 12, no. 2, pp. 600–626, 2026, doi: 10.24843/JKB.2022.v12.i02.p14.
- [3] Muhamad and I Putu Gede Arya Kusdyana, "Conservation and Utilization: Community-Based Natural and Cultural Tourism Site Design in West Bali National Park," *Jurnal Kajian Bali (Journal of Bali Studies)*, vol. 15, no. 1, pp. 92–118, 2026, doi: 10.24843/JKB.2025.v15.i01.p04.
- [4] D. Mahendra, "The Impact of Tourism on the Preservation and Transformation of Cultural Identity in Bali, Indonesia," *Studies in Social Science & Humanities*, vol. 3, no. 6, pp. 34–41, Jun. 2024, doi: 10.56397/SSSH.2024.06.05.
- [5] J. Choe, M. O' Regan, and M. A. Di Giovine, "Wellness for Tourists? Well...Mess for Locals: Socio-Culturally Sustainable Development of Wellness Tourism in Bali," *J. Dev. Stud.*, pp. 1–23, Dec. 2025, doi: 10.1080/00220388.2025.2577316.
- [6] Ni Made Arshiniwati and Ida Bagus Gede Surya Peradantha, "Reviving the Sanghyang Grodog Ritual Dance: Promoting Social Harmony and Sustainable Cultural Tourism in Nusa Lembongan Island, Bali," *Jurnal Kajian Bali (Journal of Bali Studies)*, vol. 15, no. 1, pp. 61–91, 2026, doi: 10.24843/JKB.2025.v15.i01.p03.
- [7] S. A. Suifa, D. Setiawati, M. Nur Afiat, and S. Rijal, "The Influence of Tourism Infrastructure and Online Promotion Of Tourist Visit Intention," *International Journal of Business, Law, and Education*, vol. 5, no. 2, pp. 1567–1576, May 2024, doi: 10.56442/ijble.v5i2.595.
- [8] I. Komang and B. Segara, "Destination Management in achieving Sustainable Tourism in Bali," *Bali Tourism Journal (BTJ) 2025*, vol. 9, no. 1, pp. 6–10, 2025, doi: 10.36675/btj.v9i1.122.
- [9] R. Januar, "Implementation and challenges of sustainable tourism programs in Bali," *Bali Tourism Journal*, vol. 8, no. 2, pp. 33–37, Aug. 2024, doi: 10.36675/btj.v8i2.109.
- [10] R. Salam, A. A. Samudra, A. Suradika, and I. M. Sumada, "Future policy implementation to balance growth and sustainability: Managing the impact of mass tourism in Bali," *Journal of Infrastructure Policy and Development*, vol. 8, no. 9, p. 7927, Sep. 2024, doi: 10.24294/jipd.v8i9.7927.
- [11] E. Elizabeth, H. Prabowo, A. Bandur, and R. Setiowati, "Harnessing Green Dynamic Capabilities for Sustainable Tourism Performance: The Mediating Role of Green Service Innovation in Bali's Tour and Travel SMEs," *Tourism and Hospitality*, vol. 6, no. 3, p. 156, Aug. 2025, doi: 10.3390/tourhosp6030156.
- [12] A. H. Suasapha, I. W. Suardana, P. S. Yanthy, I. N. Sunarta, and A. Ngelambong, "Selecting the Sustainable Tourist Market for Bali Based on Carbon Footprint," *Jurnal Kajian Bali*, vol. 15, no. 1, pp. 119–145, Apr. 2025, doi: 10.24843/JKB.2025.v15.i01.p05.
- [13] I. K. Surata, I. M. Sumartana, and I. G. B. R. Utama, "The impact of cultural tourism on local traditions," *JPPi (Jurnal Penelitian Pendidikan Indonesia)*, vol. 10, no. 4, pp. 672–683, Dec. 2024, doi: 10.29210/020244780.
- [14] N. P. P. Wulandari, "Multidimensional Challenges Toward Sustainable Development Tourism in Bali," *Bali Tourism Journal*, vol. 8, no. 3, pp. 56–60, Dec. 2024, doi: 10.36675/btj.v8i3.116.
- [15] I. W. Suyadnya, R. Cahyadi, L. A. Tirtayani, and A. F. Nuryani, "Overtourism sensitivity and tourism development in Canggu, Bali," *Cogent Soc. Sci.*, vol. 11, no. 1, pp. 1–21, Dec. 2025, doi: 10.1080/23311886.2025.2521364.
- [16] I. M. W. Candranegara, I. W. Mirta, and K. A. F. Putra, "Implementasi Program 'We Love Bali' Berbasis CHSE (Clean, Health, Safety, Environment) dalam Pemulihan Pariwisata Bali," *Journal of Contemporary Public Administration (JCPA)*, vol. 1, no. 1, pp. 27–32, 2021, doi: 10.22225/jcpa.1.1.2021.27-32.
- [17] I. G. P. B. S. Mananda, I. M. K. Negara, Y. Kristianto, I. G. K. H. Angligan, and C. Deuchar, "From Overtourism to Regeneration: A Penta-Helix Governance Model for Sustainable Tourism in Bali," *Tourism and Hospitality*, vol. 6, no. 5, p. 240, Nov. 2025, doi: 10.3390/tourhosp6050240.
- [18] Md. A. Islam, S. C. Goldar, S. Al Imran, Md. Halimuzzaman, and S. Hasan, "AI-Driven green marketing strategies for eco-friendly tourism businesses," *International Journal of Tourism and Hotel Management*, vol. 7, no. 1, pp. 56–60, Jan. 2025, doi: 10.22271/27069583.2025.v7i1a.125.
- [19] Y. Shen, L. Yang, and B. Deng, "NLP-Driven Sentiment Analysis for Stakeholder Feedback in Sustainable Projects," in *2024 6th International Conference on Frontier Technologies of Information and Computer (ICFTIC)*, IEEE, Dec. 2024, pp. 713–716, doi: 10.1109/ICFTIC64248.2024.10913250.
- [20] Ni Made Ary Widiastini, I Ketut Sida Arsa, Made Aristia Prayudi, and Ni Luh Putu Agustini Karta, "Strategic Collaborations and Diversification: Pathways to Sustainable Growth in Bali's Tourism Village and Business Ecosystems," *Jurnal Kajian Bali (Journal of Bali Studies)*, vol. 15, no. 1, pp. 146–172, Jan. 2026, doi: 10.24843/JKB.2025.v15.i01.p06.
- [21] W. Ardani, N. Wayan Ari Sudiartini, and N. Made Yudhaningsih, "An integrated approach for the development of a sustainable tourism village in Bali Province, Indonesia," *Problems and Perspectives in Management*, vol. 23, no. 2, pp. 667–678, Jun. 2025, doi: 10.21511/ppm.23(2).2025.48.
- [22] C. P. Gupta and V. V. R. Kumar, "A Study on Sustainable Tourism and Application of Sentiment Analysis in the Tourism Industry," in *2024 International Conference on Emerging Techniques in Computational Intelligence (ICETCI)*, IEEE, Aug. 2024, pp. 52–57, doi: 10.1109/ICETCI62771.2024.10704180.
- [23] D. Erdoğan et al., "Developing a Deep Learning-Based Sentiment Analysis System of Hotel Customer Reviews for Sustainable Tourism," *Sustainability*, vol. 17, no. 13, p. 5756, Jun. 2025, doi: 10.3390/su17135756.
- [24] X. Huang and S. Chelliah, "Attributes Influencing Tourist Satisfaction: Sentiment Analysis and Topic Modeling of Online Reviews," *Journal of China Tourism Research*, vol. 21, no. 3, pp. 819–838, Sep. 2025, doi: 10.1080/19388160.2024.2440323.
- [25] T. Ali, B. Omar, and K. Soulaimane, "Analyzing tourism reviews using an LDA topic-based sentiment analysis approach," *MethodsX*, vol. 9, p. 101894, 2022, doi: 10.1016/j.mex.2022.101894.
- [26] Y. S. Balcioglu, "Exploring consumer engagement and satisfaction in health and wellness tourism through text-mining," *Kybernetes*, vol. 54, no. 11, pp. 6886–6910, Nov. 2025, doi: 10.1108/K-12-2023-2721.
- [27] Z. Wang, P. Udomwong, J. Fu, and P. Onpium, "Destination image analysis and marketing strategies in emerging panda tourism: a cross-cultural perspective," *Cogent Business & Management*, vol. 11, no. 1, pp. 1–24, 2024, doi: 10.1080/23311975.2024.2364837.
- [28] S. Kapoor, S. Gulati, S. Verma, A. Pandey, and D. K. Vishwakarma, "Multimodal Architecture for Sentiment Recognition via employing Multiple Modalities," in *2024 2nd International Conference on Advancement in Computation & Computer Technologies (InCACCT)*, IEEE, May 2024, pp. 435–439, doi: 10.1109/InCACCT61598.2024.10551131.
- [29] S. J. Issac Sam, K. M. Jasim, and M. Babu, "Customers' metaverse service encounter perceptions: sentiment analysis and topic



- modeling,” *Journal of Hospitality Marketing & Management*, vol. 34, no. 1, pp. 92–114, Jan. 2025, doi: 10.1080/19368623.2024.2383455.
- [30] J. Li, D. Li, C. Xiong, and S. Hoi, “BLIP: Bootstrapping Language-Image Pre-training for Unified Vision-Language Understanding and Generation,” Feb. 2022, [Online]. Available: <http://arxiv.org/abs/2201.12086>
- [31] M. Grootendorst, “BERTopic: Neural topic modeling with a class-based TF-IDF procedure,” Mar. 2022, [Online]. Available: <http://arxiv.org/abs/2203.05794>

