The Image of Tourist Attraction in Bali Based on Big Data Analytics and Sentiment Analysis

Ni Wayan Sumartini Saraswati
Dept. of Engineering Science
Udayana University
Denpasar, Indonesia
0000-0002-4679-8535

I Ketut Gede Darma Putra
Dept. of Information Technology
Udayana University
Denpasar, Indonesia

Made Sudarma
Dept. of Engineering Science
Udayana University
Denpasar, Indonesia

I Made Sukarsa
Dept. of Information Technology
Udayana University
Denpasar, Indonesia

***Abstract*— Tourism development is currently very rapid due to advances in information technology. The growth of various digital tourism platforms makes it easy to obtain tourism information. Almost every time tourists leave traces of their travel experiences on digital platforms, one of them is a review. Until now, the review is only greatly beneficial to prospective tourists in terms of supporting decision-making about the consumption of tourism products and services to minimize the risk of failure. This research analyzed large amounts of reviews, known as big data, to benefit stakeholders. This research aimed to describe the image of Bali tourism that originated from tourist attraction reviews on TripAdvisor. Sentiment analysis was carried out using the Vader Lexicon method to obtain the image clarified using term frequency, bigrams, and topic-based trigrams. The analysis obtained several positive images of tourist attractions in Bali, including beautiful beaches, amazing temples, and friendly locals. Meanwhile, we recommend improvements to several negative images that we found such as dirty tourist objects, plastic waste, and disturbances from hawkers in order to get a better image of Bali tourism.**

Keywords— Tourist Attraction, Bali, Big Data, Sentiment Analysis

I. INTRODUCTION

The tourism industry has the opportunity to become one of the mainstays of the country’s largest foreign exchange earner. Growth in the tourism sector will be followed by growth in other industrial sectors, such as the transportation, craft, and creative industries. Therefore, the government's move to advance the development of the Indonesian tourism industry is considered the right strategy [1].

Currently, people are in an era where the development of the digital age is accelerating. Two factors have contributed to the rapid development of the digital era: the increasingly affordable cost of hardware and the internet. It is also supported by the development of smartphones, where people exchange information easily and quickly worldwide. The development of social media and other digital platforms also supports the formation of massive amounts of data, known as big data.

The tourism industry has also felt the impact of the rapid development of the digital world. The travel process has a new habit of consumers making online reservations for tourism products and services, such as hotels, restaurants, and tourist attraction entrance tickets. In the online reservation process, first, tourists research tourism products that will be consumed through information on the internet. One of them is online reviews. This research was conducted to minimize failures by utilizing previous user experiences with the consumption of these tourism services [2]. This information is known as electronic Word of Mouth (e-WoM), where tourists can get a clearer picture of the tourism products to be consumed from this e-WoM information.

Until now, it has only described the benefits of online reviews for potential tourists. The next question is whether this online review can also be utilized by other stakeholders in developing tourism services for better services. Thus, it requires analyzing this online review to find meaningful information that can increase the selling value of these tourist products.

Big data analytics are needed where online reviews are available in large numbers on digital platforms. The benefit of using big data is that the data extracted is based on actual consumer actions and not on data obtained from surveys, predictions, or projections. Therefore, the results are objective, not judgments that can be interpreted differently [3]. In other words, insight and information from big data have the advantage of providing an overview of the data closer to the facts.

A good tourism image is not only the basis for customers to choose tourist destinations but also an essential basis for tourism destinations to maintain their source of tourists [4]. This research developed big data analytics to study the image of tourist objects in Bali using lexicon-based sentiment analysis to support the fact above. Sentiment analysis is the process of analyzing digital text to automatically determine whether the emotional tone of the message is positive, negative, or neutral. The research data source was taken from the TripAdvisor site. It is expected that with the insights gained from processing big data analytics, an overview of the image of tourist objects in Bali can be obtained, both in the form of positive and negative images. A good understanding of this matter is expected to enable support tourism objects in

Bali to be better managed. Another important thing that can be obtained is to understand the characteristics or views of Bali. Therefore, the strategies taken by tourism object managers in Bali have become more effective and on target.

Sentiment analysis is conducted to classify a review as either a positive or negative review automatically. The next step is to observe the frequency of the words that appear most often in both positive and negative reviews with the hope that the frequency of the words that appear most often is a consumer tendency of the image of the tourist attraction.

II. Related Work

Sentiment analysis has been widely used to understand the polarity of consumer reviews related to the tourism domain. Several studies have discussed the use of machine learning to carry out sentiment analysis. Sentiment analysis on Google reviews of Borobudur and Prambanan temples was carried out by Dian Arianto [4] using a machine learning method. The aspects studied were Attractions, Amenities, Accessibility, Image, Price, and Human Resources. Machine learning methods studied were Random Forest (RF), Naïve Bayes (NB), Logistic Regression (LR), Decision Tree (DT), and Extra Tree (ET). A study by [5] aimed to mine the visitors' perceptions of Indonesia's ten most visited sites. Emotions and topics discussed in comments were two features to be extracted. Results showed that Joy was the most prominent emotion accompanying visitors' experiences. Topic modeling showed several important keywords related to preferences.

The following studies have reported the use of topic modeling other than machine learning. This study [6] obtained the polarity of sentiment and intention classes of English-language tweets related to tourism attraction in Bangkok, Chiang Mai, and Phuket. Furthermore, this study investigated the accuracy of some machine learning algorithms (Decision Tree, Random Forest, and Support Vector Machine) in predicting the polarity of sentiments and intentions of the tweets.

In a subsequent preliminary qualitative content analysis, the top ten words found in each sentiment and intention class were gathered to provide insights and suggestions to help increase tourism in Thailand. This study's methodology [7] employed a combination of topic modeling and sentiment analysis to extract valuable insights about Marrakech city from TripAdvisor reviews. Through this technique, tourism practitioners and field specialists might dive deeper into user-generated online content, leveraging aspect-based sentiment analysis to explore each destination's weaknesses and strengths.

The BERT and LSTM neural network-based methods were reported in the following studies. This study proposed an aspect-based sentiment analysis model by extracting aspect categories and corresponding sentiment polarities from tourists' reviews based on the Bidirectional Encoder Representation from Transformers (BERT) model [8]. Research using BERT was also carried out by [9], and the main purpose of this paper was to analyze a tourist destination using sentiment analysis techniques with data from Twitter and Instagram to find the most representative entities (or places) and perceptions (or aspects) of the users. Researchers studied the sentiment features of tourist online the outside community about tourist attractions in Bali in the form of advantages and disadvantages of tourist attractions in reviews from the technical perspective of natural language processing. Thus, this study proposed an improved Long Short-Term Memory (LSTM) framework for sentiment feature extraction from travel reviews [10].

Only one study reported the use of the lexicon method in analyzing reviews. Research [11] aimed to explore feedback from local and foreign tourists who had visited Labuan Bajo. In addition, it also identified the most popular tourist destinations in the area. Research data was collected from Instagram using the hashtag "*labuanbajo*". The Vader lexicon-based sentiment analysis method was used to measure sentiment polarity. The results revealed that tourist destinations in Labuan Bajo tended to evoke positive sentiments, and popular destinations frequently visited by tourists include Komodo Island, Padar Island, and Pink B.

After the polarity review, the review could be used to describe the image of a tourist object. The following are some studies on the image of tourist objects using text and data mining. Research on the image of tourist objects in Hong Kong was carried out by [12]. The key findings demonstrated that Chinese tourists have an optimistic image of Hong Kong; mainland tourists were of the view that Hong Kong's weather was both hot and stuffy, Hong Kong houses pleasant and convenient yet small hotels, together with authentic yet expensive food, convenient shopping with a good night view, friendly locals, clean and lively atmosphere and also poor service. Research by [13] analyzed a uniquely inclusive model covering text mining and data mining of destination images, reviews on tourist activities, weather forecasts, and recent events via social media to efficiently generate more user-centric and location-based thematic recommendations. Researchers used the machine learning method to get personalized augmented recommendations according to the unique preferences of travelers. This paper [14] explored the multi-dimensional label system construction for drawing destination images with the help of big data mining and fine-grained sentiment analysis technology and conducted fine-grained sentiment analysis on texts from different sources using a sentiment dictionary.

The following studies carried out further research to study tourist preferences and recommend tourism objects. Research by [15] developed a model that constructed an empirical typology of international cultural tourists who have visited Macau based on their reviews on TripAdvisor. This model could automatically indicate the probability of a visitor being of a particular cultural tourist type. This paper [16] proposed a method for supporting travel planning and sightseeing tours focusing on image search to recommend highly satisfied tourist spots based on sentiment analysis by computing image and sentiment scores.

III. Method

Sentiment analysis, also known as opinion mining, is a text mining method closely related to the Natural Processing Language (NLP) method, in which sentiment analysis extracts the expression of emotion, opinion, or mood in a text. Two technical approaches were used in sentiment analysis: machine learning and lexicon-based approaches [17]. Both approaches are shown in Figure 1 below.

Based on the two approaches above, the one using machine learning was an approach that automatically classifies text or reviews that require training data [18].

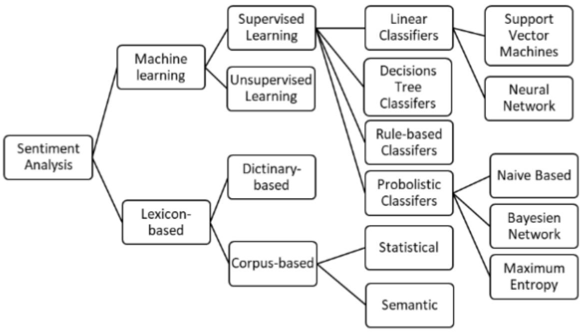

Fig. 1. Two Approaches to Conducting Sentiment Analysis [17]

A lexicon is a set of features for each word, such as word classification and emotion. It is a sentiment analysis method based on comparing the words used in the text with the previously prepared lexicon [19]. In other words, in this lexicon approach, each word is given a special weight according to the polarity of the word it has. In sentiment analysis, the words in the text will be compared with the prepared lexicon.

The lexicon approach calculated the sentiment orientation of all documents or sets of sentences from the semantic orientation of the lexicon. The semantics of this orientation can be positive, negative, or neutral [20]. A sentiment lexicon is a lexical repository consisting of sentiment terms along with their classes and scores. This sentiment lexicon plays an important role in developing sentiment analysis systems. Because each sentiment term is assigned a class and a corresponding score, it will help in computing scores at various levels, such as the level of words, sentences, or documents [21].

Lexicon does not require training data and only relies on dictionaries. In addition, Lexicon is quick to complete the analysis and provides a simple positive or negative word count [22]. This research used the lexicon approach, considering that it was ineffective for researchers to prepare training data with manual labeling given the large amount of data in this research. Thus, the lexicon method was more appropriate than the machine learning method based on the consideration.

The data used in this research came from TripAdvisor for 159 tourist attractions in Bali. After cleaning the data and grouping them based on tourist object categories, for example, beaches, temples, parks, and so on, as shown in the research flowchart in Figure 2 below.

Furthermore, sentiment analysis was carried out using the Vader Lexicon to polarize the review. The result of Vader Lexicon process, there were 5.128 negative reviews and 41.668 positive reviews regarding tourism objects in Bali. After the review sentences are polarized, we need a description of the image of each collection of positive sentiments and negative sentiments. We propose analyzing the frequency of 2 word pairs (bigrams) and the frequency of 3 word pairs (trigrams) that appear from each group.

Bigrams are sequences of two adjacent elements of a series of tokens, usually letters, syllables, or words. In contrast, trigrams are sequences of three adjacent elements of a series of tokens. This research used frequency bigrams and trigrams to get a further picture of the positive and negative images of tourist objects in Bali.

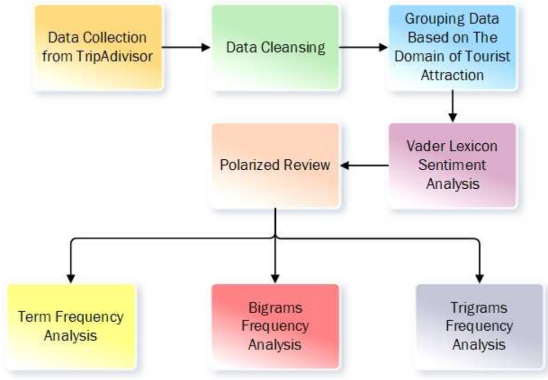

Fig. 2. Research Flowchart

Compared with other research with lexicon methods [11], we carried out further analysis to obtain images of tourist attractions using bigrams and trigrams. We found a novel method that has not been used by researchers [12]–[14] in analyzing the image of tourist attractions, that with a simple method such as determining topic words in trigrams (topic-based trigrams) we get a more focused picture of the image of tourist attractions in Bali.

IV. Experiment

This research described the frequencies of the sentiment polarity groups in the form of WordClouds, as shown in Figure 3 and Figure 4.

Fig. 3. Positive Term Frequency

Figures 3 and 4 show that several dominant words appear in both the positive and negative sentiment groups. Some words like "beach, place, people, and temple". This research also described bigrams in WordClouds, as shown in Figures 5 and 6 below. It obtained prominent word combinations, including Kuta Beach, Nusa Dua Beach, and clean beach in the positive sentiment group and dirty beach in the negative sentiment group.

In the positive group, it was found that tourist attractions in Bali were great places and worth visiting. At the same time, the research received comments regarding entrance tickets to tourist attractions and the presence of hawkers in the negative review group.

Fig. 4. Negative Term Frequency

Fig. 5. Positive Bigrams Frequency

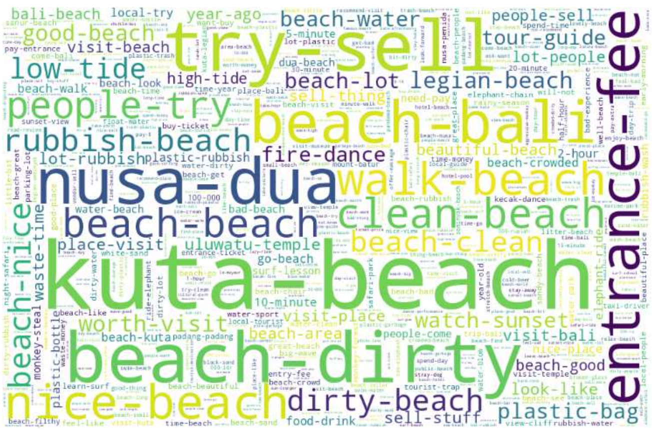

Fig. 6. Negative Bigrams Frequency

Furthermore, the research took the 50 trigrams with the greatest frequency in the words "beach, place, people, and temple" from the two review polarities, as stated in [23] to understand further. Table 1 describes the frequency of trigrams as a whole.

The trigrams data above shows some insights about the image of tourist attractions in Bali from the topic of popular tourist comments. Some popular tourist attractions that had a positive image included Nusa Dua Beach, Padang-Padang Beach, Bali Safari and Marine Park, Uluwatu Temple, Ujung Water Palace Park, Kuta Beach, Museum Blanco, Bali Bird Park, Tirta Gangga Park, Garuda Wisnu Kencana, and Blue Point Beach.

TABLE I MOST 50 TRIGRAMS

Positive Review		Negative Review	
Frequently Used Word	**Count**	**Frequently Used Word**	**Count**
('nusa', 'dua', 'beach')	533	('people', 'try', 'sell')	60
('people', 'try', 'sell')	240	('nusa', 'dua', 'beach')	44
('padang', 'padang', 'beach')	228	('try', 'sell', 'stuff')	42
('elephant', 'safari', 'park')	208	('try', 'sell', 'thing')	36
('definitely', 'worth', 'visit')	204	('local', 'try', 'sell')	35
('good', 'beach', 'bali')	160	('pay', 'entrance', 'fee')	27
('safari', 'marine', 'park')	160	('beach', 'nusa', 'dua')	25
('white', 'sand', 'beach')	157	('padang', 'padang', 'beach')	23
('beach', 'nusa', 'dua')	145	('visit', 'kuta', 'beach')	22
('try', 'sell', 'thing')	142	('waste', 'time', 'money')	20
('bali', 'safari', 'marine')	142	('lot', 'rubbish', 'beach')	18
('try', 'sell', 'stuff')	128	('beach', 'dirty', 'lot')	17
('beach', 'white', 'sand')	120	('kuta', 'beach', 'dirty')	17
('place', 'visit', 'bali')	119	('people', 'sell', 'thing')	17
('beach', 'kuta', 'beach')	115	('beach', 'kuta', 'beach')	17
('pay', 'entrance', 'fee')	112	('kecak', 'fire', 'dance')	16
('great', 'place', 'visit')	111	('rubbish', 'beach', 'water')	15
('staff', 'friendly', 'helpful')	111	('sell', 'thing', 'beach')	14
('nice', 'place', 'visit')	109	('take', '2', 'hour')	13
('kecak', 'fire', 'dance')	108	('hawker', 'try', 'sell')	13
('visit', 'uluwatu', 'temple')	107	('people', 'sell', 'stuff')	12
('white', 'sandy', 'beach')	106	('white', 'sandy', 'beach')	12
('highly', 'recommend', 'visit')	103	('good', 'beach', 'bali')	12
('crystal', 'clear', 'water')	102	('visit', 'uluwatu', 'temple')	12
('good', 'water', 'park')	99	('plastic', 'bag', 'bottle')	12
('visit', 'kuta', 'beach')	96	('kuta', 'legian', 'seminyak')	11
('beach', 'good', 'beach')	95	('nice', 'place', 'visit')	11
('ujung', 'water', 'palace')	93	('monkey', 'forest', 'ubud')	11
('visit', 'place', 'bali')	92	('kuta', 'beach', 'nice')	11
('good', 'place', 'visit')	91	('famous', 'beach', 'bali')	11

('garuda', 'wisnu', 'kencana')	90	('safari', 'marine', 'park')	11
('kuta', 'beach', 'good')	88	('place', 'visit', 'bali')	10
('stay', 'nusa', 'dua')	85	('beach', 'plastic', 'rubbish')	10
('kuta', 'beach', 'great')	85	('well', 'beach', 'bali')	10
('local', 'try', 'sell')	84	('beach', 'good', 'beach')	10
('nice', 'clean', 'beach')	83	('need', 'pay', 'entrance')	10
('beach', 'good', 'place')	83	('beach', 'lot', 'rubbish')	10
('beach', 'great', 'place')	80	('try', 'sell', 'ware')	10
('watch', 'sun', 'set')	78	('kuta', 'beach', 'beach')	10
('legian', 'beach', 'hotel')	77	('kuta', 'beach', 'time')	10
('beach', 'nice', 'beach')	76	('go', 'kuta', 'beach')	10
('good', 'time', 'visit')	76	('famous', 'kuta', 'beach')	10
('beach', 'beautiful', 'beach')	76	('come', 'kuta', 'beach')	10
('spend', 'couple', 'hour')	74	('bali', 'safari', 'marine')	10
('watch', 'kecak', 'dance')	74	('come', 'try', 'sell')	9
('beach', 'watch', 'sunset')	73	('plastic', 'bottle', 'bag')	9
('kecak', 'dance', 'performance')	73	('dirty', 'beach', 'lot')	9
('water', 'crystal', 'clear')	72	('go', 'early', 'morning')	9
('beautiful', 'place', 'visit')	72	('visit', 'nusa', 'dua')	9
('don', 'antonio', 'blanco')	72	('blue', 'bird', 'taxi')	9

Some interesting things that gave a positive image of tourist attractions in Bali were beaches with white sand, crystal clear water, sunset experiences, friendly and helpful tour operators, Kecak Fire Dance performances, great places to relax, a low and affordable entrance fee, and the fact that the temple was located on a cliff.

On the other hand, the negative image obtained from trigrams was that there were some negative comments about Nusa Dua Beach, Padang-Padang Beach, Kuta Beach, Uluwatu Temple, and Bali Safari and Marine Park, with a much smaller number than positive reviews. The other tourist attractions mentioned were the Ubud Monkey Forest and Mount Batur. Some things defined as giving a bad image included dirty beaches, beaches full of trash, hawkers, plastic waste, paying for an entrance ticket, leaving early in the morning, receding beaches, crowded beaches, dirty beach water, and monkeys who steal stuff. Suppose it was associated with the condition of the tourist attraction. In that case, the possibility of monkey disturbance was experienced at Uluwatu temple or monkey forest Ubud, while having to leave very early in the morning could be associated with climbing Mount Batur or watching dolphin at Lovina Beach.

We recognize the limitations of presenting research results in this manuscript. The complete trigrams for the keywords beach, temple, people, beautiful and rubbish can be downloaded at [23]

Based on the results of these studies, we obtained positive images of beaches in Bali as beaches with white sand, clean beaches, beaches with beautiful sunsets, beaches with many visitors, long stretches of beach, famous beaches, many nice beaches, great beaches for surfing, clear beach water, walks on Kuta Beach, and beaches for water sports. Some beaches mentioned in positive reviews included Nusa Dua Beach, Padang-Padang Beach, Kuta Beach, Blue Point Beach, Legian Beach, and Pandawa Beach. As for the negative image of beaches in Bali, this research obtained the following popular comments. There were beaches with a lot of garbage, plastic waste on the beach, a very dirty beach, a low tide beach, a bustling and dirty beach, a small and bustling beach, a bustling Kuta beach, hawkers on the beach, beach entrance tickets, rainy season on the beach, and beaches covered in trash. The beaches mentioned in the negative review included Nusa Dua Beach, Padang-Padang Beach, and Kuta Beach.

Some popular temples from the review with positive images include the Uluwatu Temple, Taman Ayun Temple, Ulun Danu Temple, Gunung Kawi Temple, Pasar Agung Temple, Tanah Lot Temple, and Tirta Empul Temple. Some interesting things about the temple in a positive image included the location of the temple on the edge of a cliff, the Kecak dance at the temple, the sunset at Uluwatu Temple, the temple from the 11th century, the amazing view of the temple, the beautiful temple in Bali, the temple with rice fields, the temple with a view of the beach, and the experience of using a sarong to enter the temple. Meanwhile, this research obtained a negative image of the temple as follows. The temples mentioned included the Uluwatu Temple, Gunung Kawi Temple, Ulun Danu Temple, Danau Buyan Temple, Tanah Lot Temple, and Taman Ayun Temple. Some interesting things related to negative images are as follows. There was a fire dance at the temple, a big and complex temple, damaged artifacts, a small temple door, cloudy weather at the temple, payments at the temple, and aggressive monkeys at the temple.

The positive image regarding the keyword "people" is about the *acung* traders (not all have a negative opinion about the existence of the *acung* traders), busy visitors on the beach, lots of visitors surfing, friendly local residents, people who offer massage services, friendly and helpful people, kind-hearted locals, visitors enjoying surfing, visitors enjoying their time, watching visitors swim, visitors visiting temples, and visitors taking photos. While the negative image was related to hawkers, crowded visitors at the beach, people who throw garbage, people with dirty beaches, people who are constantly rude, rude hawkers, drunken beachgoers, queues of people, people selling tickets, beggars, groups of people filming events, many people are frenzied, and local community corruption.

We tried to collect positive trigrams information with the keyword beautiful to understand positive images and perceptions about beauty in Bali. Some of the information obtained related to beautiful beaches, beautiful places to visit, beautiful Kuta beaches, beaches with sand, beautiful white water, beautiful sunsets, clean and beautiful beaches, beautiful places worth visiting, beautiful water palace, beautiful sea view, beautiful Legian Beach, beautiful Tirta

Gangga temple, beautiful garden, beautiful clear water, beautiful Kecak Dance, beautiful Padang-Padang Beach, beautiful water park, beautiful Nusa Dua, beautiful temple, beautiful cliff temple, beautiful small beach, beautiful place of relaxation, beautiful terraced rice fields, beautiful animals, and beautiful Pandawa Beach.

This research also chose the keyword "rubbish" to describe the image of the waste problem in Bali. It obtained information related to a lot of trash on the beach, water on the beach with trash, plastic waste on the beach, lots of trash and dirty, little trash, people throwing trash, trash on the beach, trash straws and plastic bags on the beach, embarrassingly huge amounts of plastic waste, burying pits, swimming with a lot of trash, trash with local tourists, raw waste trash, dead fish trash, leaving trash in rocks, and paying attention to plastic waste.

V. Conclusion

Generally, the image of tourist attractions in Bali was positive based on TripAdvisor reviews. Some popular tourist objects that received a positive image included Nusa Dua Beach, Padang-Padang Beach, Kuta Beach, Blue Point Beach, Legian Beach, Pandawa Beach, Uluwatu Temple, Taman Ayun Temple, Ulun Danu Temple, Gunung Kawi Temple, Pasar Agung Temple, Tanah Lot Temple, and Tirta Empul Temple. Tourists highlighted beautiful, clean beaches, friendly locals, beautiful sunset views, and stunning temples. On the other hand, trash at tourist objects was a bad image for tourism, as there were hawkers who rudely offered their wares, complaints about entrance tickets to tourist attractions, and disturbance of monkeys at several tourist objects. Experiences in tourism that received positive images included surfing, watching the Kecak dance, and elephant shows at the Bali Safari Park.

Acknowledgment

Our thanks go to Udayana University for facilitating this research well.

References

[1] Biro Hukum dan Komunikasi Publik, "Pariwisata Kini Jadi Andalan Pendulang Devisa Negara," *Kemenpar.Go.Id*, 2015. http://www.kemenpar.go.id/asp/detil.asp?c=16&id=2959.

[2] F. Aprilia and A. Kusumawati, "Influence of Electronic Word of Mouth on Visitor's Interest to Tourism Destinations," *J. Asian Financ. Econ. Bus.*, vol. 8, no. 2, pp. 993–1003, 2021, doi: 10.13106/jafeb.2021.vol8.no2.0993.

[3] N. Kraus, "The Big Data Revolution in Tourism," *Tourism Review*, 2017. https://www.tourism-review.com/big-data-technology-used-in-the-tourism-sector-news10293 (accessed Apr. 28, 2022).

[4] D. Arianto and I. Budi, "Aspect-based Sentiment Analysis on Indonesia's Tourism Destinations Based on Google Maps User Code-Mixed Reviews (Study Case: Borobudur and Prambanan Temples)," *Proc. 34th Pacific Asia Conf. Lang. Inf. Comput.*, vol. 2019, pp. 359–367, 2020, [Online]. Available: https://aclanthology.org/2020.paclic-1.41.

[5] H. Irawan, G. Akmalia, and R. A. Masrury, "Mining tourist's perception toward Indonesia tourism destination using sentiment analysis and topic modelling," in *ACM International Conference Proceeding Series*, 2019, no. 1, pp. 7–12, doi: 10.1145/3361821.3361829.

[6] N. Leelawat *et al.*, "Twitter data sentiment analysis of tourism in Thailand during the COVID-19 pandemic using machine learning," *Heliyon*, vol. 8, no. 10, p. e10894, 2022, doi: https://doi.org/10.1016/j.heliyon.2022.e10894.

[7] 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: https://doi.org/10.1016/j.mex.2022.101894.

[8] M. Chu, Y. Chen, L. Yang, and J. Wang, "Language interpretation in travel guidance platform: Text mining and sentiment analysis of TripAdvisor reviews," *Front. Psychol.*, vol. 13, 2022, doi: 10.3389/fpsyg.2022.1029945.

[9] M. Santiago, Viñán-Ludeña, and L. M. de Campos, "Discovering a tourism destination with social media data: BERT-based sentiment analysis," *J. Hosp. Tour. Technol.*, vol. 13, no. 5, pp. 907–921, 2022, doi: 10.1108/JHTT-09-2021-0259.

[10] M. Fu and L. Pan, "Sentiment Analysis of Tourist Scenic Spots Internet Comments Based on LSTM," *Math. Probl. Eng.*, vol. 2022, 2022, doi: 10.1155/2022/5944954.

[11] J. Setiawan, V. Gousander, and I. Prasetiawan, "Unmasking the Sentiments of Labuan Bajo: An Instagram-based Analysis for Tourism Insights through VADER Sentiment Analysis," *G-Tech J. Teknol. Terap.*, vol. 7, no. 3, pp. 967–976, 2022.

[12] Q. Jiang, C. S. Chan, S. Eichelberger, H. Ma, and B. Pikkemaat, "Sentiment analysis of online destination image of Hong Kong held by mainland Chinese tourists," *Curr. Issues Tour.*, vol. 24, no. 17, pp. 2501–2522, 2021, doi: 10.1080/13683500.2021.1874312.

[13] S. P. R. Asaithambi, R. Venkatraman, and S. Venkatraman, "A Thematic Travel Recommendation System Using an Augmented Big Data Analytical Model," *Technologies*, vol. 11, no. 1, 2023, doi: 10.3390/technologies11010028.

[14] Y. Luo, L. Deng, L. Xiao, B. Wu, J. Pan, and J. Wang, "Research on destination images with information containing sentiment classification driven by multi-source data," in *2022 4th International Conference on Data-driven Optimization of Complex Systems (DOCS)*, 2022, pp. 1–6, doi: 10.1109/DOCS55193.2022.9967750.

[15] S. Qi, C. U. I. Wong, N. Chen, J. Rong, and J. Du, "Profiling Macau cultural tourists by using user-generated content from online social media," *Inf. Technol. Tour.*, vol. 20, no. 1–4, pp. 217–236, 2018, doi: 10.1007/s40558-018-0120-0.

[16] K. Oooka and Y. Wang, "A Tourist Spot Recommendation Method Based on Image Search and Sentiment Analysis," in *2022 IEEE 11th Global Conference on Consumer Electronics (GCCE)*, 2022, pp. 670–674, doi: 10.1109/GCCE56475.2022.10014382.

[17] I. A. Ozen, "Tourism Products and Sentiment Analysis," in *Advances In Hospitality And Tourism Information Technology*, vol. 18, no. 9781732127586, University of South Florida (USF) M3 Publishing, 2021, p. 2.

[18] V. Bonta, N. Kumaresh, and N. Janardhan, "A Comprehensive Study on Lexicon Based Approaches for Sentiment Analysis," *Asian J. Comput. Sci. Technol.* , vol. 8, no. S2, pp. 1–6, Mar. 2019, doi: 10.51983/AJCST-2019.8.S2.2037.

[19] M. Al-Shabi, "Evaluating the performance of the most important Lexicons used to Sentiment analysis and opinions Mining," *Int. J. Comput. Sci. Netw. Secur.*, vol. 20, no. 1, pp. 51–57, 2020.

[20] N. Gupta and R. Agrawal, "Application and techniques of opinion mining," *Hybrid Comput. Intell.*, pp. 1–23, Jan. 2020, doi: 10.1016/B978-0-12-818699-2.00001-9.

[21] A. Khattak, M. Z. Asghar, A. Saeed, I. A. Hameed, S. Asif Hassan, and S. Ahmad, "A survey on sentiment analysis in Urdu: A resource-poor language," *Egypt. Informatics J.*, vol. 22, no. 1, pp. 53–74, Mar. 2021, doi: 10.1016/J.EIJ.2020.04.003.

[22] Z. Drus and H. Khalid, "Sentiment Analysis in Social Media and Its Application: Systematic Literature Review," *Procedia Comput. Sci.*, vol. 161, pp. 707–714, Jan. 2019, doi: 10.1016/J.PROCS.2019.11.174.

[23] Ni Wayan Sumartini Saraswati, "Trigrams result of Bali Tourist Attraction Image," *Kaggle*, 2023. https://www.kaggle.com/datasets/sumartinisaraswati/trigrams-result (accessed Jul. 25, 2023).