Harnessing Social Media Data for Sentiment Analysis of Tourist Attractions in Trat Province, Thailand using the Random Forest Machine Learning Approach

NARONG PLEERUX[1], PHANNIPHA ANURUKSAKORNKUL[2],
PARADORN BOONPOR[3], PARINYA NAKPATHOM[4]

[1]Faculty of Geoinfomatics, Burapha University,
Chon Buri, THAILAND

[2]Faculty of Humanities and Social Sciences, Burapha University,
Chon Buri, THAILAND

[3]Department of National Parks, Wildlife and Plant Conservation,
Ministry of Natural Resources and Environment,
Bangkok, THAILAND

[4]Internatioal Collage, Burapha University,
Chon Buri, THAILAND

Abstract: Tourism and service industries are vital economic drivers worldwide, and social media platforms play a pivotal role in disseminating and gathering tourist reviews. This study employed the random forest algorithm to analyze tourist reviews of attractions in Trat Province, Thailand, using data collected from the Tripadvisor website between 2014 and 2023. From the results, key issues impacting these destinations were identified and categorized into four main areas, i.e., scenery, facilities, safety, and accessibility. With a high accuracy rate of 99.65%, the analysis revealed that 98.66% of the reviews reflected positive sentiment, underscoring the province's appeal. However, the findings of this study also highlight critical challenges, particularly in terms of facilities and safety, which require attention to realize sustainable tourism management. The findings provide valuable insights for stakeholders to enhance the quality of tourism services in Trat, aligning with the province's aspirations to elevate its status to a primary tourist destination in Thailand.

Keywords: natural language processing, sentiment analysis, random forest, social media, tourism

Received: April 9, 2024. Revised: September 15, 2024. Accepted: November 16, 2024. Published: December 31, 2024.

1. Introduction

The tourism and service sectors are vital sources of income for nearly all nations globally [1]. As tourism has grown, there has been a corresponding increase in tourism-related websites and social networks, which has resulted in a substantial increase in the creation and sharing of tourism-related information and opinions [2]. Many people express their views and seek information via various social media platforms [3], e.g., X (formerly Twitter), Facebook, and Tripadvisor [4–5]. The global popularity of social media platforms continues to grow, playing an increasingly integral role in people's lives across all dimensions [6], and such platforms have become the primary mode of communication for people worldwide [7].

Information shared on social media appears in various formats, including text, images, and videos, with most content reflecting opinions on events or topics, e.g., reviews of tourist attractions on Tripadvisor [8–9] or photo sharing on Instagram [10]. Social media platforms enable users to express their opinions openly and directly. Social media data offers real-time information and is cost-effective compared to traditional survey methods [6]. In addition, social media data can cover extensive time periods and allow for the selection of specific time frames; thus, social media data are ideal for studying the development or behavior of individuals and groups over different periods [11].

Data from social media are utilized to examine people's satisfaction, attitudes, and emotions toward products, services, or different issues through sentiment analysis. Such analyses are typically categorized into three levels,. i.e., positive, negative, and neutral [12–13]. Sentiment analysis is widely applicable in various fields, particularly in tourism and service industries, and it is frequently used to analyze customer opinions about restaurants [14–15], hotels [16–19], and tourist destinations [20–21].

The primary objective of this study is to analyze tourist feedback about attractions in Trat Province, Thailand, using machine learning models and identify key issues affecting these destinations. According to the Tourism Authority of Thailand, Trat is classified as a secondary tourist province in Thailand. A secondary tourist province is defined as one that receives fewer than four million tourists per year [22]. The Thai government has been making efforts to elevate these provinces to primary tourist destinations by implementing various promotional measures, e.g., tax incentives [23] and organizing events to attract more visitors. Consequently, tourist feedback from social media platforms like Tripadvisor is a crucial source of information to manage and plan tourism effectively in Trat Province according to current needs and circumstances.

The remainder of this paper is organized as follows. Section 2 presents the target data and methodology used in this study. Section 3 discusses the research findings and corresponding analysis. Finally, the paper is concluded in Section 4.

2. Data and Methodology

2.1 Study Area

Trat Province, located at the easternmost tip of Thailand, is shaped like an elephant's head and spans 2,819 km2. It is the fourth smallest province in the eastern region of the country and ranks 56th in size nationwide. The province is 315 km from Bangkok.

Trat is known for its rich biodiversity, featuring waterfalls, mountains, the sea, and beautiful coral reefs, as well as abundant natural resources. Trat Province also has a rich history and is home to internationally famous tourist spots, e.g., Koh Chang and Koh Kood (Koh means island). In addition, there are several notable community-based tourism sites, e.g., Ban Nam Chiao and Ban Tha Ranae. Trat is also an important area for fruit cultivation in Thailand.

2.2 Data Collection

In this study, tourist reviews of attractions in Trat Province from 2014 to 2023 were gathered from Tripadvisor. The data collection process was performed by web scraping using Python with the Selenium and Beautiful Soup libraries [24].

The acquired dataset comprised a total of 8,492 Tripadvisor reviews, each containing the name of the attraction, the date of the review, and the review content. Note that only reviews written in English were considered in this study.

2.3 Data Preprocessing

After collecting the tourist reviews of attractions in Trat Province, the data were preprocessed using natural language processing techniques to clean the text. Here, the first step involved converting all text to lowercase English. Then, irrelevant characters, e.g., punctuation, URLs, symbols, numbers, and special characters, were removed. The text was then tokenized into smaller linguistic units. Finally, a list of stop words was applied to filter out insignificant words, including prepositions, conjunctions, pronouns, classifiers, and emojis (e.g., "I," "you," "we," "me," "the," and "is").

2.4 Sentiment Analysis

The sentiment analysis in this study involves five essential steps: review labeling, review splitting, which is vital for preparing data for training and testing the machine learning model, text representation, model development for sentiment analysis, and model performance evaluation. The specifics of these steps are outlined below.

1) The dataset of 8,492 tourist reviews of attractions in Trat Province was used for the review labeling process. Here, a random sample of 10% of the reviews (849 reviews) was selected, and then three experts in tourism and data science categorized the reviews into positive, neutral, and negative sentiment groups. The labeled data were then utilized to train a machine learning model.

2) Review splitting is an essential process to prepare data to train and test a machine learning model. The review dataset was divided into three subsets, i.e., a training set, a validation set, and a test set with respective proportions of 64%, 16%, and 20%. The model utilized these data to learn and identify patterns and relationships. The validation set was used to evaluate and refine the model throughout the training process, and the test set was used to assess the model's performance after training and validation. Importantly, the test set is completely distinct from both the training and validation sets, which ensures that the model does not encounter any data from the test set.

3) The text representation process is a critical step in preparing inherently nonnumerical text. Converting text into numerical data is essential to realize effective analysis. This study utilized the term frequency-inverse document frequency vectorizer (TFIDFVectorizer) as the text representation method. This technique considers both TF and the significance of words across different documents (i.e., the IDF) [25].

4) The random forest (RF) algorithm was employed for sentiment analysis. Introduced by Breiman in 2001 [26], the RF algorithm integrates the principles of random subspaces and bagging. The decision tree forest algorithm is trained on multiple decision trees, each of which uses slightly different subsets of the data [27]. The RF method is adept at handling complex and diverse datasets, and it effectively mitigates overfitting problems.

The compound score for each review generated by the model ranges from −1 to 1 and is categorized into three groups, i.e., positive (≥0.05), neutral (≥−0.05 and <0.05), and negative (<−0.05) [28].

5) The evaluation of the model's performance includes a comprehensive overview of its effectiveness, incorporating several metrics, e.g., accuracy, precision, recall, and F1-score [29].

In sentiment analysis model evaluation, accuracy, precision, recall, and F1-score are crucial metrics for measuring performance. While accuracy reflects the proportion of correctly classified sentiments overall, it can be misleading in imbalanced datasets. Precision highlights the correctness of predicted positive sentiments, and recall shows how well the model identifies actual positives. Since these metrics can trade off, the F1-score combines them into a single value, offering a more balanced measure, especially when dealing with imbalanced data or when both false positives and false negatives have different impacts. Together, these metrics offer a well-rounded evaluation of model performance in sentiment classification.

2.5 Identification of Issues

This section categorizes the issues identified at the target tourist attractions in Trat Province based on the negative review from tourists, which were analyzed using the RF model. The issues were grouped into four categories, i.e., scenery (representing the beauty and cleanliness of the attractions), facilities (the adequacy of amenities, staff, and services), safety (crime and the safety of the services provided at the attractions), and accessibility (representing ease of use, affordability, and the availability of multiple options to access attractions).

3. Results and Discussion

The findings are discussed in terms of three perspectives, i.e., the number of tourist reviews, the results of analyzing the tourist reviews of attractions using the RF model, and the categorization of the identified issues at the attractions based on negative reviews.

3.1 Number of Tourist Reviews

The analysis of tourist comments of attractions in Trat Province, which were collected from the Tripadvisor website from 2014 to 2023, revealed a total of 8,492 comments. The peak year for comments was 2016, with 1,628 comments, which was followed by 2017 with 1,321 comments and 2019 with 1,215 comments. However, during the COVID-19 pandemic, there was a considerable decline in the number of comments, with only 133 recorded in 2021. This decrease was largely due to the government's lockdown measures and international travel bans implemented to mitigate the spread of the virus [30], thereby resulting in a significant drop in travel and commentary. As the pandemic situation improved, the number of comments began to increase steadily, as shown in Fig. 1.

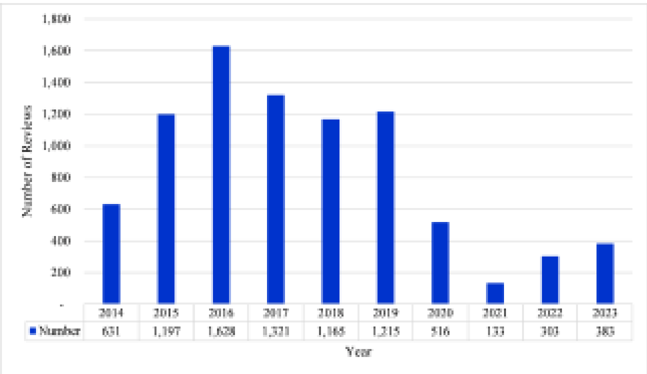

Fig. 1. Annual number of tourist reviews on Tripadvisor from 2014 to 2023.

An analysis of the top 10 attractions with the most reviews revealed that six are natural attractions, and the remaining four are diving agencies. For example, Klong Plu Waterfall was found to be the most reviewed site, with 858 reviews (representing 10.10% of the total reviews). Following this, White Sand Beach was reviewed 697 times (representing 8.21% of the total reviews), and the Scuba Dawgs diving school ranked third with 454 (5.36%). Note that most of the top 10 reviewed attractions are natural attractions. This can be attributed to Trat Province's location along Thailand's eastern coast, which boasts stunning and famous beaches and islands, e.g., Koh Chang, where Klong Plu Waterfall, Khlong Prao Beach, and Bang Bao Beach are located. Other attractions, e.g., Koh Kood and Koh Mak, also draw significant tourist interest, resulting in numerous reviews of these attractions. In addition, four of the top 10 sites are diving agencies, reflecting the province's clear waters, diverse marine life, and colorful coral reefs, which make it a popular destination for snorkeling and scuba diving. Both Thai and international tourists frequently visit dive sites, e.g., Koh Rang, Koh Chang, and Blueberry Hill, as shown in Table I.

TABLE I. TOP 10 ATTRACTIONS IN TRAT PROVINCE WITH THE HIGHEST NUMBER OF TRIPADVISOR REVIEWS

Ranking	Attractions	Frequency	%[a]
1st	Klong Plu Waterfall	858	10.10
2nd	White Sand Beach	697	8.21
3rd	Scuba Dawgs	455	5.36
4th	BB Divers	388	4.57
5th	Khlong Prao Beach	329	3.87
6th	BB Divers Koh Kood	297	3.50
7th	Koh Kood Divers	289	3.40
8th	Bang Bao Beach	246	2.90
9th	Lonely Beach	240	2.83
10th	Kai Bae Beach	238	2.80

a. Percentage of 8,492 reviews

3.2 Sentiment of Reviews

The examination of the tourist reviews of attractions in Trat Province indicates that the model achieved an overall accuracy of 99.65%. Remarkably, 98.66% of the reviews (totaling 8,378 reviews) conveyed positive sentiments toward the attractions in the region. In comparison, neutral reviews accounted for 0.73% (62 reviews), and negative review comprised only 52 reviews (representing 0.61%), as shown in Table II.

TABLE II. NUMBER OF POSITIVE, NEUTRAL, AND NEGATIVE TOURIST REVIEWS OF ATTRACTIONS IN TRAT PROVINCE OBTAINED BY THE RF MODEL

Year	Positive	Neutral	Negative	Total
2014	623	6	2	631
2015	1178	9	10	1,197
2016	1596	22	10	1,628
2017	1313	3	5	1,321
2018	1153	8	4	1,165
2019	1189	10	16	1,215
2020	513	1	2	516
2021	131	0	2	133
2022	303	0	0	303
2023	379	3	1	383
Total	**8,378**	**62**	**52**	**8,492**
%	**98.66**	**0.73**	**0.61**	**100.00**

After performing a detailed analysis of the 8,492 reviews, we found that they pertained to 189 distinct tourist attractions and service establishments. Most tourists expressed positive feedback about various aspects of these attractions. For example, Klong Plu Waterfall was praised for its beauty and easy of accessibility, and White Sand Beach received comments highlighting its long stretch of white sandy shore and clear, clean waters, thereby making it an ideal destination for families with children. In addition, some tourists noted that White Sand Beach has numerous street food vendors and excellent restaurants.

In contrast, several issues were identified when analyzing the negative feedback about these attractions, including cleanliness problems, service quality concerns, and incidents of animal cruelty. Specific issues at White Sand Beach included problems related to the beach's condition and cleanliness, safety hazards, and the availability, quality, and ethical standards of accommodations and services. For example, at elephant camps, tourists reported an unwelcoming atmosphere, poor service, and concerns about animal abuse and ethical practices. Even Klong Plu Waterfall, which received the most positive feedback, as mentioned previously, was not without its issues. The problems cited included management inefficiencies, cost and value concerns, and safety issues, particularly regarding the safety of children.

3.3 Issues Identification from Negative Reviews

To gather information about the identified issues, complaints, and suggestions from tourists regarding the attractions in Trat Province, we analyzed and categorized 52 negative reviews into four main areas, i.e., scenery, facilities, safety, and accessibility. Note that a single comment may address multiple issues; thus, we counted the total number of distinct issues described in each comment.

Among the negative reviews, the identified issues were related to 25 different attractions and services, with a total of 93 mentions across the four identified categories. The most frequently mentioned issue was facilities, which was referenced 37 times, accounting for 39.78% of the total. This was followed by safety, which was mentioned 26 times (27.96%), and scenery, which appeared 19 times (20.43%). We found that accessibility was the least discussed issue, with only 11 mentions (representing 11.83%), as shown in Table III.

TABLE III. NUMBER AND PERCENTAGE OF ISSUES IDENTIFIED AT TOURIST ATTRACTIONS IN TRAT PROVINCE DERIVED FROM NEGATIVE REVIEWS ANALYZED USING THE RF MODEL

Issues	Natural attractions		Man-made attractions		Total	
	No.	%	No.	%	No.	%
Scenery	12	32.43	7	12.51	19	20.43
Facilities	10	27.03	27	48.21	37	39.78
Safety	8	21.62	18	32.14	26	27.96
Accessibility	7	18.92	4	7.14	11	11.83
Total	**37**	**100.00**	**56**	**100.00**	**93**	**100.00**

The 25 tourist attractions mentioned previously can be classified into two main categories, i.e., natural attractions and man-made attractions, which helps facilitate the discussion of various issues. Focusing on natural attractions, tourists provided 20 negative reviews (out of a total of 52) that were associated with seven different natural sites. Among these sites, White Sand Beach received the highest number of negative comments, with seven in total. This was followed by Klong Plu Waterfall and Lonely Beach, each of which received four negative reviews. The remaining sites, i.e., Bang Bao Beach, Wai Chaek Beach, Than Mayom Waterfall, Ao Noi Beach, and Kai Bae Beach, were each mentioned only once in a negative context.

Among the negative reviews related to natural attractions, tourists most frequently mentioned concerns about scenery, accounting for 32.43% of the feedback. This was followed by issues regarding facilities (27.03%), safety (21.62%), and accessibility (18.92%). Regarding scenery, tourists primarily criticized the cleanliness, especially at several beaches, e.g., White Sand Beach, Lonely Beach, and Kai Bae Beach, where problems with litter and trash were identified.

In terms of facilities, tourists expressed dissatisfaction with the entrance fees at Klong Plu Waterfall and Than Mayom Waterfall, which were considered too high (200 and 100 baht or approximately 5.40 and 2.70 US dollar for foreign adults and children, respectively). Note that these waterfalls are located within a national park, and the fees are set according to the park's regulations.

Concerning safety and accessibility, tourists were particularly concerned about the steep roads on Koh Chang, given the island's hilly terrain interspersed with flat areas, which could contribute to accidents. In addition, concerns were raised about the overall quality of the roads.

For man-made attractions, tourists provided 32 negative reviews out of a total of 52. The most frequently criticized aspect was the facilities, accounting for 48.21% of the negative feedback, followed by safety concerns at 32.14%, and scenery at 12.51%. Accessibility received the fewest negative mentions, at only 7.14%. When analyzing the issues at these man-made attractions, the majority were found to be related to tour and diving agencies, restaurants, pubs and bars, and animal camps.

Regarding issues with facilities, many tourists reported negative experiences, e.g., poor service and a lack of responsibility from staff at diving agencies. In restaurants, pubs, and bars, complaints included rude and unprofessional behavior from staff, and poor cleanliness and sanitation on buses and boats. Another significant concern was related to animal camps, particularly elephant camps, which are prevalent in Trat Province. Tourists expressed ethical concerns about riding elephants, the abuse and distress of the animals, and their overall poor treatment.

The next most significant issue at man-made attractions and services was safety, with several key concerns identified, including food hygiene and safety in restaurants, as well as dangerous diving practices and instructor attitudes impacting safety at tour and diving agencies. In terms of scenery, tourists mentioned an unwelcoming atmosphere and dirty locations. The final issue, i.e., accessibility, was primarily focused on ferry services, with complaints about overcrowded ferries and chaotic transfers.

3.4 Theoretical Implications

This study employed the RF algorithm to analyze tourist reviews of attractions in Trat Province using review data acquired from the Tripadvisor website. The RF algorithm is particularly advantageous when handling large volumes of unstructured textual reviews [31]. It combines high accuracy with relatively quick training times; thus, the RF algorithm is ideal for complex sentiment analysis that requires both precision and efficiency [32]. The RF algorithm can manage both regression and classification tasks with a high degree of accuracy and a reduced risk of overfitting [33]. Its key advantage lies in the high accuracy of its results, which has resulted in its widespread use in sentiment analysis, as exhibited by the following accuracy rates: 99.04% [32], 82.91% [34], 86.00% [35], and 83.50% [36]. In the current study, the RF algorithm achieved an accuracy of 99.65%, further underscoring its effectiveness in terms of applying machine learning technology to analyze sentiment from social media, particularly in the context of tourism and service-related feedback.

3.5 Managerial Implications

This study performed an in-depth analysis of the negative tourist feedback, categorizing it into specific issues to identify problems faced at tourist attractions in Trat Province, as directly reported by the visitors. The insights gained from this analysis can inform strategies to address and managing these challenges. For example, the most frequently mentioned issue involved facilities, with complaints focusing on negative service experiences, e.g., impolite and unaccommodating staff, as well as cleanliness concerns in restrooms, boats, and buses. Safety was the second most significant concern, divided into two key areas, i.e., food hygiene and the safety of services, particularly in activities like diving. To address these concerns, businesses, e.g., restaurants, ferry operators, and tour and diving companies, should establish, implement, and maintain clear service standards and provide staff effective training on both customer etiquette and diving safety.

Issues related to the natural scenery were primarily observed at nature-based attractions, where the main problems were related to cleanliness and litter on the beaches. The litter was traced back to two major sources, i.e., marine debris, particularly during the monsoon season when large amounts of trash are washed ashore, and waste left by both tourists and locals. To address this, relevant authorities, e.g., the local municipal government, should develop and implement comprehensive plans to manage cleanliness, including providing adequate trash bins and increasing the frequency of beach cleanups.

Effectively and sustainably resolving these four key issues requires collaborative efforts from all stakeholders, including tourism business operators, government agencies, local residents, and tourists. Only through such cooperation can these challenges be addressed successfully.

4. Conclusion

In this study, the sentiment analysis technique was used to evaluate the satisfaction and emotions of tourists or customers regarding tourism and services in Trat Province, Thailand. This study employed the RF algorithm, enhanced by machine learning, to analyze tourist reviews of attractions using data acquired from the Tripadvisor website. The results demonstrate that a significant majority of tourists, approximately 98.66%, expressed positive views about the attractions in the target region. This high level of positive sentiment indicates that the attractions are both beautiful and appealing, successfully captivating and impressing tourists. Nevertheless, this positive feedback can also guide further improvements and enhancements to the province's tourism offerings. Conversely, the negative feedback can be categorized into several key issues, with facilities being the top priority, followed by safety, scenery, and accessibility issues. These insights, including comments, criticisms, and suggestions, can be relayed to relevant authorities to aid in managing, planning, and addressing the challenges at tourist attractions in Trat Province to better fulfill the needs of tourists.

This study has two key limitations. First, the RF algorithm was employed to analyze social media feedback, and highly accurate results were obtained; however, future research should consider incorporating additional models, e.g., naïve Bayes, support vector machine, gradient boosting, and other models, to optimize the outcomes. Second, this study relied exclusively on reviews from the Tripadvisor website, which may lead to potential bias. Therefore, to achieve more comprehensive and balanced results, future research should include data from other social media platforms, e.g., X (formerly Twitter), and Facebook.

Acknowledgment

This research project was financed by the Fundamental Fund 2024, Burapha University.

References

[1] S. Ainin, A. Feizollaha, N.B. Anuarb, and N.A. Abdullah, "Sentiment analyses of multilingual tweets on halal tourism," Tour. Manag. Perspect., vol. 34, pp. 100658, April 2020.

[2] M. Ghane'i‑Ostad, H. Vahdat‑Nejad, and M. Abdolrazzagh‑Nezhad, "Detecting overlapping communities in LBSNs by fuzzy subtractive clustering," Soc. Netw. Anal. Min., vol. 8, pp. 23, March 2018.

[3] D. Arora, K.F. Li, and S.W. Neville, "Consumers' sentiment analysis of popular phone brands and operating system preference using twitter data: A feasibility study," The 2015 IEEE 29th International Conference on Advanced Information Networking and Applications, pp. 680–686, March 2015.

[4] N. Leelawat, S. Jariyapongpaiboon, A. Promjun, S. Boonyarak, K. Saengtabtim, A. Laosunthara, A.K. Yudha, and J. Tang, "Twitter data sentiment analysis of tourism in Thailand during the COVID-19 pandemic using machine learning," Heliyon, vol. 8, pp. e10894, October 2022.

[5] E. Polyzos, A. Fotiadis, and T.C. Huan, "The asymmetric impact of Twitter sentiment and emotions: Impulse response analysis on European tourism firms using micro-data," Tour. Manag., vol. 104, pp. 104909, October 2024.

[6] Y. Cui, and Q. He, Travel behavior analysis with social media data and smartphone GPS data. Tech Report, CA: University of Buffao, 2019.

[7] N.B. Lassen, R. Madsen, and R. Vatrapu, "Predicting iphone sales from iphone tweets," The 2014 IEEE 18th International Enterprise Distributed Object Computing Conference, pp. 81–90, September 2014.

[8] A. Valdivia, M.V. Luzon, and F. Herrera, "Sentiment analysis in TripAdvisor," IEEE Intell. Syst., vol. 32, pp. 72–77, August 2017.

[9] N. Pleerux, and A. Nardkulpat, "Sentiment analysis of restaurant customer satisfaction during COVID-19 pandemic in Pattaya, Thailand," Heliyon, vol. 9, pp. e22193, November 2023.

[10] T. Bastrygina, W.M. Lim, R. Jopp, and M.A. Weissmann, "Unraveling the power of social media influencers: Qualitative insights into the role of Instagram influencers in the hospitality and tourism industry," J. Hosp. Tour. Manag., vol 58, pp. 214–243, March 2024.

[11] M. Picornell, T. Ruiz, M. Lenormand, J.J. Ramasco, T. Dubernet, and E. Frias-Martines, "Exploring the potential of phone call data to characterize the relationship between social network and travel behavior," Transportation, vol. 42, pp. 647–668, July 2015.

[12] A. Valdivia, M.V. Luzon, and F. Herrera, "Sentiment analysis on TripAdvisor: are there inconsistencies in user reviews?," "The

12th International Conference on Hybrid Artificial Intelligent Systems," pp. 15–25, June 2017.

[13] D. Obembe, O. Kolade, F. Obembe, A. Owoseni, and O. Mafimisebi, "Covid-19 and the tourism industry: An early stage sentiment analysis of the impact of social media and stakeholder communication," Int. J. Inf. Manag. Data Insights, vol. 1, pp. 100040, November 2021.

[14] H. Zhang, P. Ji, J. Wang, and X. Chen, "A novel decision support model for satisfactory restaurants utilizing social information: a case study of TripAdvisor.com," Tour. Manag., vol. 59, pp. 281–297, April 2017.

[15] P. Pleerux, W. Boonpook. A. Chaiyo, and A. Nardkulpat, "TripAdvisor restaurant customer reviews analysis in Bangkok, Thailand," The 2023 International Conference on Cyber Management and Engineering, pp. 536–539, January 2023.

[16] C. Tsai, K. Chen, Y. Hu, and W. Chen, "Improving text summarization of online hotel reviews with review helpfulness and sentiment," Tour. Manag., vol. 80, pp. 104122, October 2020.

[17] Y. Chang, C. Ku, and C. Chen, "Using deep learning and visual analytics to explore hotel reviews and responses," Tour. Manag., vol. 80, pp. 104129, October 2020.

[18] J. Zhang, X. Lu, and D. Liu, "Deriving customer preferences for hotels based on aspect-level sentiment analysis of online reviews," Electron. Commer. Res. Appl., vol. 49, pp. 101094, September–October 2021.

[19] Y. Bian, R. Ye, J. Zhang, and X. Yan, "Customer preference identification from hotel online reviews: A neural network based fine-grained sentiment analysis," Comput. Ind. Eng., vol. 172, pp. 108648, October 2022.

[20] N.G. Uğur, and A. Akbıyık, "Impacts of COVID-19 on global tourism industry: A cross-regional comparison," Tour. Manag. Perspect., vol. 36, pp. 100744, October 2020.

[21] Z. Abbasi-Moud, H. Vahdat-Nejad, and J. Sadri, "Tourism recommendation system based on semantic clustering and sentiment analysis," Expert Syst. Appl., vol. 167, pp. 114324, April 2021.

[22] TAT Intelligence Center, (2018). "The compass of Tourism 4/2018," [online] Available at:<https://www.dla.go.th/upload/news/type9/2018/11/40628_1.pdf ?time=1544920032593>

[23] Ratchakitcha, (2019). "Ministerial Regulation No. 344 B.E. 2562 (2019) issued under the Revenue Code regarding tax exemption Volume 136 Part 65 A," [Online] Available at: <https://www.rd.go.th/fileadmin/user_upload/kormor/newlaw/mr344.pdf>

[24] S. Han, and C.K. Anderson, "Web Scraping for Hospitality Research: Overview, Opportunities, and Implications," Cornell Hosp. Q., vol. 62, pp. 89–104, July 2020.

[25] M. Arora, V. Mittal, and P. Aggarwal, "Enactment of tf-idf and word2vec on text categorization," The 3rd International Conference on Computing Informatics and Networks, pp. 199–209, March 2019.

[26] L. Breiman, "Random Forests," Mach. Learn., vol. 45, pp. 5–32, October 2001.

[27] Y. Al Amrani, M. Lazaar, and K.E. El Kariri, "Random forest and support vector machine based hybrid approach to sentiment analysis," Procedia Comput. Sci., vol. 127, pp. 511–520, 2018.

[28] A.R. Pai, M. Prince, and C.V. Prasannakumar, "Real-time Twitter sentiment analytics and visualization using Vader", The 2nd International Conference on Intelligent Technologies, pp. 1–4, June 2022.

[29] O. Rainio, J. Teuho and R. Klén, "Evaluation metrics and statistical tests for machine learning," Sci. Rep., vol. 14, pp. 6086, July 2024.

[30] Ratchakitcha, (2021). "Regulation issued under section 9 of the emergency decree on public administration in emergency situations B.E. 2548 (2005) (No. 1)," [Online] Available at: <http://www.ratchakitcha.soc.go.th/DATA/PDF/2563/E/ 069/T_0010.PDF>

[31] N. Saraswathi, T.S. Rooba, and S. Chakaravarthi, "Improving the accuracy of sentiment analysis using a linguistic rule-based feature selection method in tourism reviews," Meas.: Sens., vol. 29, pp. 100888, October 2023.

[32] T. Anderson, S. Sarkar, and R. Kelley, "Analyzing public sentiment on sustainability: A comprehensive review and application of sentiment analysis techniques," J. Nat. Lang. Process., vol. 8, pp. 100097, September 2024.

[33] IBM, (2024). "What is random forest?," [Online] Available at: <https://www.ibm.com/topics/random-forest>.

[34] V.A. Fitri, R. Andreswari, M.A. Hasibuan, "Sentiment analysis of social media Twitter with case of anti-LGBT campaign in Indonesia using Naïve Bayes, Decision Tree, and Random Forest algorithm," Procedia Comput. Sci., vol. 161, pp.765–772, July 2019.

[35] M. Shahzad, C. Freeman, M. Rahimi and H. Alhoori, "Predicting Facebook sentiments towards research," J. Nat. Lang. Process., vol. 3, pp. 100010, June 2023.

[36] M. Ramzy & B. Ibrahim, "User satisfaction with Arabic COVID-19 apps: Sentiment analysis of users' reviews using machine learning techniques," Inf. Process. Manage., vol. 61, pp. 103644, May 2024.

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

The authors equally contributed in the present research, at all stages from the formulation of the problem to the final findings and solution.

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself

No funding was received for conducting this study.

Conflict of Interest

The authors have no conflicts of interest to declare that are relevant to the content of this article.