Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 411-420
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5416

Sentiment and Toxicity Analysis of Tourism-Related Video through Vader, Textblob, and Perspective Model in Communalytic

Yerik Afrianto Singgalen

Faculty of Business Administration and Communication, Tourism Study Program, Atma Jaya Catholic University of Indonesia, Jakarta, Indonesia
Email: yerik.afrianto@atmajaya.ac.id
Correspondence Author Email: yerik.afrianto@atmajaya.ac.id
Submitted: **26/06/2024**; Accepted: **30/06/2024**; Published**: 30/06/2024**

Abstract−This study leverages the Tourism and Travel Content Analysis (TTCA) framework to explore user sentiment and behavior in response to digital travel content. Utilizing sentiment analysis models such as VADER and TextBlob, the research analyzed 13,162 posts, revealing that 13.92% were negative, 15.02% neutral, and 71.06% positive, according to VADER. At the same time, TextBlob classified 10.47% as unfavorable, 26.51% as neutral, and 63.02% as positive. Additionally, toxicity scores calculated using Detoxify and Perspective models showed a range from low to high levels of toxic content, highlighting issues like identity attacks, insults, profanity, and threats. The findings underscore the effectiveness of well-crafted narratives in digital content for influencing tourist behavior and visit intentions. However, limitations were noted in the model's ability to fully capture emotional and cultural nuances. Future research should incorporate more advanced analytical tools and diverse datasets to overcome these limitations. Ultimately, the TTCA framework provides valuable insights for enhancing digital marketing strategies and improving user engagement in the tourism sector.

Keywords: Sentiment; Toxicity; Tourism; Vader; Textblob; Perspective

1. INTRODUCTION

Tourism-related videos significantly impact the branding and image of travel destinations. Content creators increasingly produce travel content as personal documentation and a marketing strategy for these destinations [1], [2]. These videos' visual appeal and narrative elements are crucial in shaping potential tourists' perceptions and preferences [3]–[7]. This phenomenon underscores destinations' need to engage in digital storytelling and visual marketing to stay competitive [8]–[13]. Therefore, tourism boards and businesses must harness the power of video content to enhance promotional efforts and attract a broader audience.

Enhancing user engagement in digital content necessitates strategic efforts to captivate public interest, encompassing elements from video titles to meticulously crafted storyboards. Compelling video titles are the initial point of attraction, compelling viewers to click and explore the content further [14]–[17]. Furthermore, well-structured storyboards ensure a coherent narrative flow, maintaining viewer interest throughout the video [18]–[21]. By incorporating these elements, digital content captures and retains audience attention, leading to higher engagement rates [22]. Consequently, a comprehensive approach to content creation, emphasizing intriguing titles and engaging narratives, is essential for maximizing user interaction.

Optimizing digital content for travel presents the challenge of aligning narratives and videos with the experiences and preferences of travelers. Understanding varied travel motivations and interests requires crafting content that resonates with diverse traveler expectations [23]. Additionally, producing high-quality, authentic videos that capture the essence of travel destinations is essential to engage viewers effectively [24]. Therefore, tailoring content to reflect genuine traveler experiences and preferences is crucial for achieving meaningful audience engagement [25]. Ultimately, overcoming these challenges enhances the effectiveness of travel content in inspiring and attracting potential tourists.

This study aims to identify viewer sentiment towards travel video content to understand perceptions of specific travel destinations. Analyzing comments and reactions allows for insights into the emotional responses elicited by these videos. Additionally, sentiment analysis provides valuable information on how different aspects of the videos influence viewers' opinions and attitudes. Understanding these sentiments is crucial for developing effective marketing strategies and improving destination branding. Ultimately, the findings from this research will offer a deeper comprehension of how travel videos shape public perception and influence travel decisions.

The urgency of this research lies in its potential to enhance the effectiveness of digital marketing strategies in the tourism industry. Rapid advancements in digital media have significantly altered how travel destinations are perceived and chosen, necessitating a deeper understanding of consumer engagement with online content [26]. By investigating the factors that drive viewer sentiment and interaction, this research provides critical insights for tailoring marketing efforts to meet audience expectations better [27]. Consequently, timely and comprehensive analysis in this area is essential for maintaining a competitive edge in the dynamic and evolving landscape of tourism marketing.

This research's theoretical and practical contributions are multifaceted, offering valuable insights into the dynamics of digital content and consumer behavior in the tourism industry. Theoretically, it advances the understanding of how digital media influences tourist perceptions and decision-making processes, providing a foundation for future academic inquiries [28]. Practically, the findings offer actionable recommendations for tourism marketers to enhance engagement and effectiveness in digital strategies [29]. This dual contribution underscores the

Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 411-420
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5416

significance of the study, bridging the gap between academic theory and real-world application to foster more effective and targeted tourism marketing practices.

A notable limitation of this research lies in the reliance on the sentiment analysis models Vader, Textblob, and Perspective API for sentiment evaluation and toxicity score calculation. While effective in many contexts, these models may not fully capture the nuances of sentiment in diverse and culturally specific expressions found in travel content. Additionally, variations in algorithmic interpretations lead to discrepancies in sentiment and toxicity assessments. Despite these constraints, the chosen models provide a foundational framework for sentiment analysis, highlighting the need for future research to develop more sophisticated tools tailored to the intricacies of digital travel content.

Similar research often focuses on the CRISP-DM methodology for sentiment analysis and toxicity scoring, calculating average to maximum values. In contrast, this study constructs a customized framework tailored to specific analysis needs, utilizing the Communalytic model for data processing. The distinctive approach of this research allows for more nuanced insights, adapting to the unique requirements of the dataset and research objectives. Consequently, the tailored framework offers a more precise and contextually relevant analysis, highlighting the importance of methodological flexibility in sentiment analysis and toxicity evaluation. This differentiation underscores the study's contribution to advancing analytical techniques in digital content research.

2. RESEARCH METHODOLOGY

2.1 Gap Analysis

The gap analysis of viewer sentiment towards tourism-related video reviews, particularly traveler documentation, reveals several topics that warrant in-depth exploration using content analysis. Identifying these gaps highlights the need to contextual examine video narratives to uncover underlying viewer perceptions and preferences. Such an approach enables a comprehensive understanding of how specific content elements influence viewer sentiment. Therefore, addressing these identified gaps through detailed content analysis is crucial for advancing knowledge in the field and enhancing the effectiveness of tourism-related digital content. This targeted analysis underscores the importance of narrative context in shaping viewer engagement and sentiment.

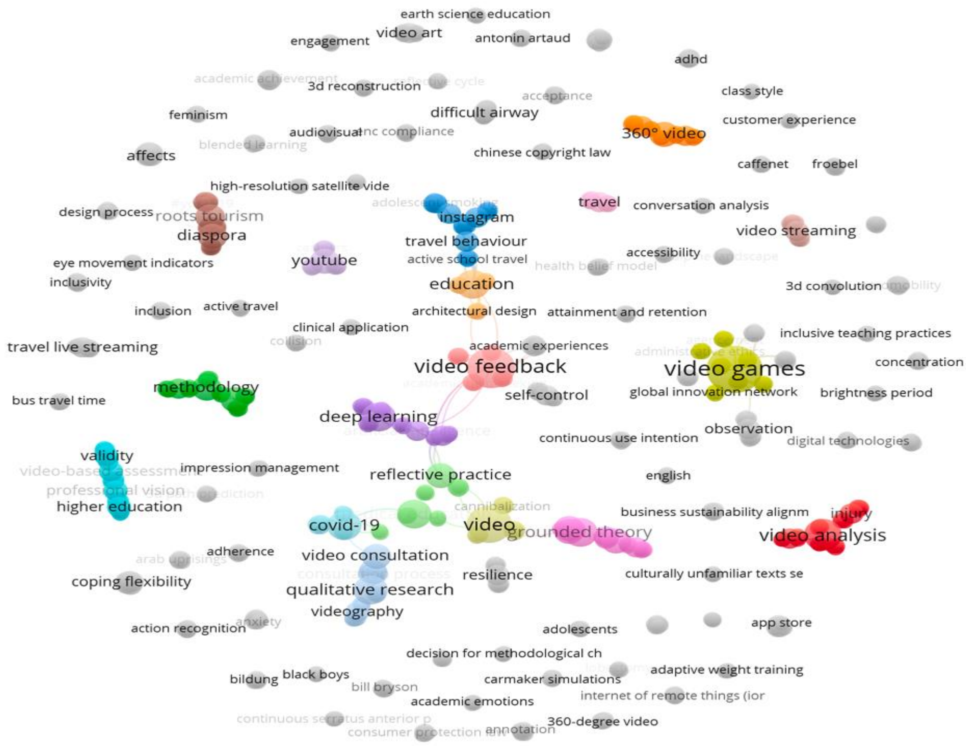

Figure 1. Network Visualization of Digital Content Topics (VosViewer)

Figure 1 shows the network visualization of digital content topics. The study of tourism digital content from the perspective of sentiment and toxicity analysis requires further development. Current research provides a foundational understanding but lacks a comprehensive exploration of the nuanced emotional responses and negative interactions the digital content elicits [30]. Enhancing these studies could lead to more sophisticated methods for evaluating and mitigating adverse viewer reactions, improving content quality and audience engagement [31]. Such advancements are essential for refining digital marketing strategies and fostering a more positive online environment

Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 411-420
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5416

for tourism promotion [32]. Therefore, expanding research in this area is crucial for the continued evolution and effectiveness of tourism digital content.

Several studies on sentiment analysis and toxicity analysis remain limited to evaluating the performance of models or algorithms. However, specific discussions on tourism video content must be analyzed based on market perceptions and preferences through destination video reviews [33]. This focused approach allows a more accurate understanding of how video content influences tourist behavior and destination appeal. Therefore, continued research is essential, focusing on outcomes that contribute theoretically and practically, enhancing the overall efficacy and relevance of sentiment and toxicity analysis in tourism.

2.2 Tourism and Travel Content Analysis (TTCA)

The data processing framework employed in analyzing reviews is the Tourism and Travel Content Analysis (TTCA) framework. TTCA is meticulously designed to align with the specific objectives of data processing, ensuring the generation of information pertinent to content creators, tourists, and destination managers. By tailoring the framework to these diverse stakeholders, TTCA facilitates the extraction of actionable insights that enhance content relevance and effectiveness. This targeted approach underscores the framework's value in addressing the distinct needs of various audiences within the tourism industry. Ultimately, TTCA contributes significantly to optimizing digital content strategies and improving destination management practices.

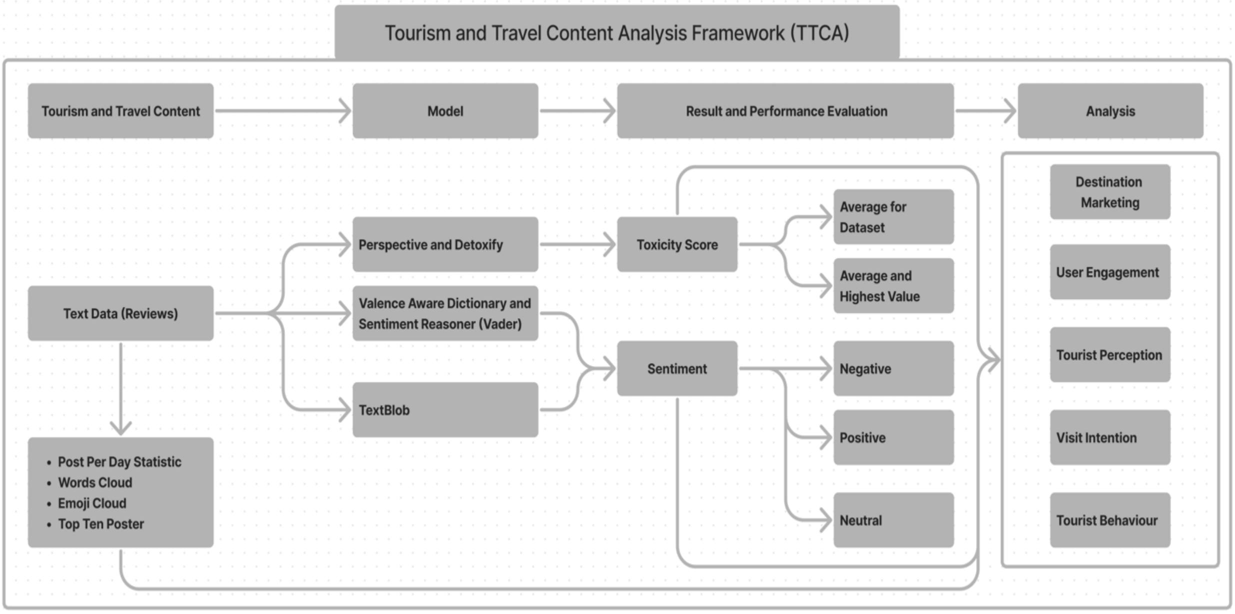

Figure 2. Tourism and Travel Content Analysis Framework

Figure 2 shows the implementation of the TTCA framework. The TTCA framework is constructed based on the performance of Communalytic, which is utilized in analyzing the effectiveness of destination marketing. It also evaluates user engagement, tourist perceptions, visit intentions, and tourist behavior. By leveraging Communalytic's robust analytical capabilities, TTCA provides comprehensive insights into various aspects of digital content and its impact on tourism. This approach ensures that marketing strategies are data-driven and tailored to meet the needs and preferences of potential tourists. Ultimately, the TTCA framework significantly enhances the understanding and optimization of digital marketing efforts in the tourism sector.

The TTCA framework boasts several advantages, including using processed data with Perspective and Detoxify models for toxicity score calculation and Vader and Textblob models for sentiment classification. Additionally, metrics such as posts-per-day, word clouds, emoji clouds, and top ten posters serve as foundational elements for analyzing user engagement. These metrics effectively represent content creators' performance in generating compelling narratives for destination marketing and reviewing travel destinations. Consequently, TTCA provides a robust analytical tool for understanding and enhancing the impact of digital content in the tourism industry, ensuring that marketing strategies are engaging and effective.

2.2.1 Text Data (Reviews)

Tourism and Travel Content must be analyzed contextually, focusing on textual data from video reviews. In this context, the video under analysis, identified by the ID T2Avd3tFHc, was published by the Yes Theory Channel with the title "Traveling to the 'Worst' Country in Europe" on August 1, 2022, and has garnered 7,313,858 views. The substantial viewership provides a rich dataset for examining viewer sentiments and perceptions. Analyzing such high-engagement content reveals critical insights into audience reactions and the effectiveness of the video's narrative. Consequently, this contextual analysis is essential for understanding the impact of tourism-related digital content.

Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 411-420
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5416

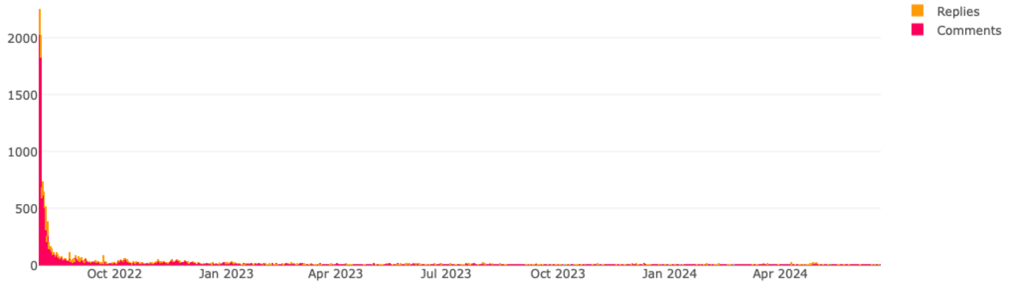

Figure 3. Post-per-day Statistic of the Video (Communalytic)

Figure 3 shows the content's post-per-day statistics. Based on the post-per-day statistics, the data reveals the following: on July 31, 2022, there were 2028 posts; on August 1, 2022, there were 1831 posts; on August 2, 2022, there were 590 posts; on August 3, 2022, there were 616 posts; on August 4, 2022, there were 506 posts; on August 5, 2022, there were 313 posts; on August 6, 2022, there were 207 posts; on August 7, 2022, there were 255 posts; on August 8, 2022, there were 148 posts; and on August 9, 2022, there were 136 posts. These statistics indicate a significant spike in activity around the release date, followed by a gradual decline in user engagement. This pattern suggests that initial interest and interaction are highest immediately following publication, emphasizing the importance of timing in digital content strategy. Understanding this trend is crucial for optimizing content release schedules to maximize user engagement and visibility.

Figure 4. Top Ten Poster (Communalytic)

Figure 4 shows the top ten posters. Based on the top ten poster data, the following statistics have been identified: @dyawr contributed 450 posts, @mr.sunshine3947 contributed 33 posts, @MostlyHarmless9 contributed 32 posts, @33Verst contributed 13 posts, @andrewrobinson2565 contributed 12 posts, @sujansilwaal4614 contributed 11 posts, @dacialogan6605 contributed ten posts, another user with details not provided contributed ten posts, @Call_Upon_YAH contributed ten posts, and @ihavenolife4675 contributed ten posts. This distribution shows a significant disparity in engagement levels among top contributors, with @dyawr demonstrating exceptionally high activity. Such concentrated contributions by a few users may indicate the presence of highly engaged or influential individuals within the audience. This insight is crucial for understanding user dynamics and leveraging key influencers in digital content strategy to enhance engagement and reach.

Figure 5. Frequently Used Words (Communalytic)

Figure 5 shows the frequently used words. Based on the frequently used words data, it observed that the most common terms include "video" (153), "guys" (149), "Mikhail" (148), and "people" (124). Additionally, words such as "grandpa" (101), "country" (100), and "Moldova" (98) are prominent, indicating critical themes in the content. The

Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 411-420
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5416

presence of words like "love" (84), "beautiful" (49), and "good" (49) suggests a generally positive sentiment associated with the video. These frequently used words provide valuable insights into the video content's primary subjects and emotional tone, highlighting its appeal and areas of viewer engagement. Consequently, understanding these patterns informs more effective content creation and marketing strategies in tourism.

Figure 6. Emoji Cloud (Communalytic)

Figure 6 shows the emoji clouds. Based on the frequently used words data, it observed that the most common terms include "video" (153), "guys" (149), "Mikhail" (148), and "people" (124). Additionally, words such as "grandpa" (101), "country" (100), and "Moldova" (98) are prominent, indicating critical themes in the content. The presence of words like "love" (84), "beautiful" (49), and "good" (49) suggests a generally positive sentiment associated with the video. These frequently used words provide valuable insights into the video content's primary subjects and emotional tone, highlighting its appeal and areas of viewer engagement. Consequently, understanding these patterns informs more effective content creation and marketing strategies in tourism.

Based on the context of tourism data, an in-depth examination through data processing is essential to yield accurate information. Practical analysis requires meticulous data handling to uncover meaningful insights that reflect trends and patterns in tourist behavior and preferences. Additionally, employing advanced analytical techniques ensures the reliability and relevance of the findings. Therefore, rigorous data processing enhances the precision of tourism-related information and supports informed decision-making for stakeholders. Ultimately, such comprehensive analysis is vital for advancing tourism research and improving strategic initiatives within the industry.

2.2.2 Model

The models utilized for calculating toxicity scores are the Detoxify and Perspective models. Detoxify is designed to identify and quantify toxic language in digital content through advanced machine-learning techniques. Similarly, the Perspective model, developed by Jigsaw and Google, assesses text for various forms of toxicity, providing a comprehensive toxicity score. Employing these models ensures a robust and accurate content evaluation, highlighting areas of concern. Therefore, integrating Detoxify and Perspective models significantly enhances the reliability of toxicity assessment in digital content analysis.

The models employed for sentiment analysis are Vader and Textblob. Vader, short for Valence Aware Dictionary and Sentiment Reasoner, is particularly effective for social media text due to its ability to capture both polarity and intensity of sentiment. Textblob, on the other hand, is a versatile tool that provides simple APIs for everyday natural language processing tasks, including sentiment classification. These models ensure a comprehensive textual sentiment analysis, capturing many emotional nuances. Consequently, Vader and Textblob significantly enhance the accuracy and depth of sentiment analysis in digital content evaluation.

2.2.3 Result and Performance Evaluation

The models utilized for calculating toxicity scores are the Detoxify and Perspective models. Detoxify is designed to identify and quantify toxic language in digital content through advanced machine-learning techniques. Similarly, the Perspective model, developed by Jigsaw and Google, assesses text for various forms of toxicity, providing a comprehensive toxicity score. Employing these models ensures a robust and accurate content evaluation, highlighting areas of concern. Therefore, integrating Detoxify and Perspective models significantly enhances the reliability of toxicity assessment in digital content analysis.

The models employed for sentiment analysis are Vader and Textblob. Vader, short for Valence Aware Dictionary and Sentiment Reasoner, is particularly effective for social media text due to its ability to capture both polarity and intensity of sentiment. Textblob, on the other hand, is a versatile tool that provides simple APIs for everyday natural language processing tasks, including sentiment classification. These models ensure a comprehensive textual sentiment analysis, capturing many emotional nuances. Consequently, Vader and Textblob significantly enhance the accuracy and depth of sentiment analysis in digital content evaluation.

2.2.4 Analysis

Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 411-420
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5416

At the analysis stage, this study links toxicity identification and sentiment classification results with the context of tourism destination marketing, user engagement, tourist perception, visit intention, and tourist behavior. The study comprehensively explains how digital content influences potential tourists by correlating toxicity and sentiment data with these key tourism metrics. Such connections highlight the impact of negative and positive sentiments on marketing effectiveness and user engagement levels. Consequently, these insights inform more targeted and effective marketing strategies. Ultimately, this integrative analysis enhances understanding of the interplay between digital content and tourism dynamics.

3. RESULT AND DISCUSSION

The discussion in this research is divided into three sections: the analysis and results of toxicity calculations, sentiment classification outcomes, and an examination of user engagement and visit intention. The toxicity analysis section delves into the severity and prevalence of harmful content identified using advanced models. The sentiment classification results provide insights into the overall emotional tone of user comments, categorized into negative, neutral, and optimistic sentiments. Finally, the discussion on user engagement and visit intention explores how these sentiments influence user interaction and potential travel decisions. This structured approach ensures a comprehensive understanding of the multifaceted impacts of digital content on tourism.

3.1 Toxicity Score and Analysis

The benefits of toxicity analysis in discussing tourism and travel content are substantial. It identifies and mitigates harmful or offensive language, enhances user experience, and maintains a positive online environment. Additionally, toxicity analysis provides insights into potential areas of user dissatisfaction or negative sentiment, allowing for targeted interventions and content improvements. This process not only fosters a safer and more welcoming digital space but also enhances the credibility and attractiveness of tourism content. Ultimately, the integration of toxicity analysis is crucial for optimizing content quality and ensuring effective audience engagement in the tourism industry.

Based on the data processing in Communalytic, the analysis of 13,845 posts (out of 15,322) using the Perspective API revealed the following results: Toxicity ranged from 0.09913 to 0.96069, Severe Toxicity from 0.00844 to 0.61261, Identity Attack from 0.03436 to 0.81821, Insult from 0.04854 to 0.81319, Profanity from 0.05673 to 0.95403, and Threat from 0.01434 to 0.70934. These findings indicate significant variability in harmful content within the dataset. The high upper ranges for metrics such as Profanity and Toxicity underscore the necessity for effective content moderation strategies. Consequently, these insights are essential for improving the quality and safety of digital tourism content. Ultimately, comprehensive toxicity analysis enhances understanding of user interactions and informs better content management practices.

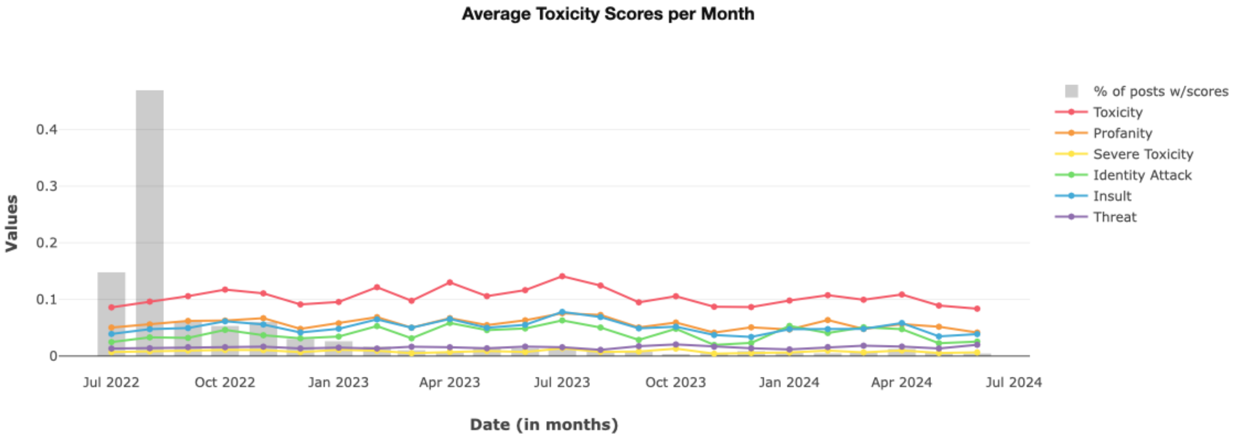

Figure 7. Toxicity Score (Communalytic)

Figure 9 shows the toxicity score based on the score per month. The toxicity scores range from 0.09913 to 0.96069, indicating a broad spectrum of potentially harmful interactions. Severe Toxicity scores, spanning from 0.00844 to 0.61261, and Identity Attack scores, ranging from 0.03436 to 0.81821, highlight the presence of extreme negativity and personal attacks. Similarly, Insult and Profanity scores, with ranges of 0.04854 to 0.81319 and 0.05673 to 0.95403, respectively, demonstrate frequent occurrences of offensive language. The Threat scores, albeit lower, from 0.01434 to 0.70934, still signify potential threats within the dataset. These results underscore the critical need for robust content moderation and intervention strategies to maintain a safe and welcoming digital environment. Consequently, this comprehensive analysis provides valuable insights into user behavior and content dynamics, essential for enhancing digital tourism content quality and safety.

Thus, toxicity analysis enhances sentiment analysis. Identifying and quantifying negative interactions provides a deeper understanding of the emotional tone within user comments. This additional layer of analysis helps distinguish

Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 411-420
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5416

between mild dissatisfaction and severe discontent, enabling more precise targeting of interventions. Consequently, integrating toxicity analysis refines the overall sentiment evaluation process, leading to more accurate insights. Ultimately, this approach improves content moderation and enhances user engagement and satisfaction by fostering a more positive online environment.

3.2 Sentiment Classification

Sentiment classification is crucial in understanding reviewers' sentiments based on an analysis of 13,411 out of 15,322 posts. This process categorizes user comments into positive, negative, and neutral sentiments, providing a clear overview of the overall emotional tone. By systematically classifying these sentiments, it becomes possible to identify prevalent attitudes and trends within the user base. Consequently, this classification aids in tailoring content and strategies to meet audience preferences and expectations better. Ultimately, sentiment classification is essential for gaining nuanced insights into user feedback and enhancing digital tourism content's effectiveness.

	# of Posts	Negative Sentiment [-1..-0.05]	Neutral Sentiment (-0.05..0.05)	Positive Sentiment [0.05..1]
VADER (English/EN)	13162	1832 (13.92%)	1977 (15.02%)	9353 (71.06%)
TextBlob (English/EN)	13162	1378 (10.47%)	3489 (26.51%)	8295 (63.02%)
TextBlob (French/FR)	109	0 (0.00%)	102 (93.58%)	7 (6.42%)
TextBlob (German/DE)	92	2 (2.17%)	81 (88.04%)	9 (9.78%)

Figure 8. Vader and TextBlob Performance in Sentiment Classification

Figure 8 shows the performance of the Vader and TextBlob models. Based on the results of sentiment classification, the following findings were obtained: for VADER (English/EN), out of 13,162 posts, 1,832 (13.92%) exhibited negative sentiment [-1..-0.05], 1,977 (15.02%) showed neutral sentiment (-0.05..0.05), and 9,353 (71.06%) displayed positive sentiment [0.05..1]. For TextBlob (English/EN), out of 13,162 posts, 1,378 (10.47%) were negative, 3,489 (26.51%) were neutral, and 8,295 (63.02%) were positive. For TextBlob (French/FR), out of 109 posts, none were harmful (0.00%), 102 (93.58%) were neutral, and 7 (6.42%) were positive. For TextBlob (German/DE), out of 92 posts, 2 (2.17%) were negative, 81 (88.04%) were neutral, and 9 (9.78%) were positive. These results indicate a predominantly positive sentiment across all languages analyzed, with English posts showing a higher proportion of positive sentiment than French and German posts. This comprehensive sentiment distribution provides valuable insights for tailoring content strategies to align with audience sentiments in different linguistic contexts.

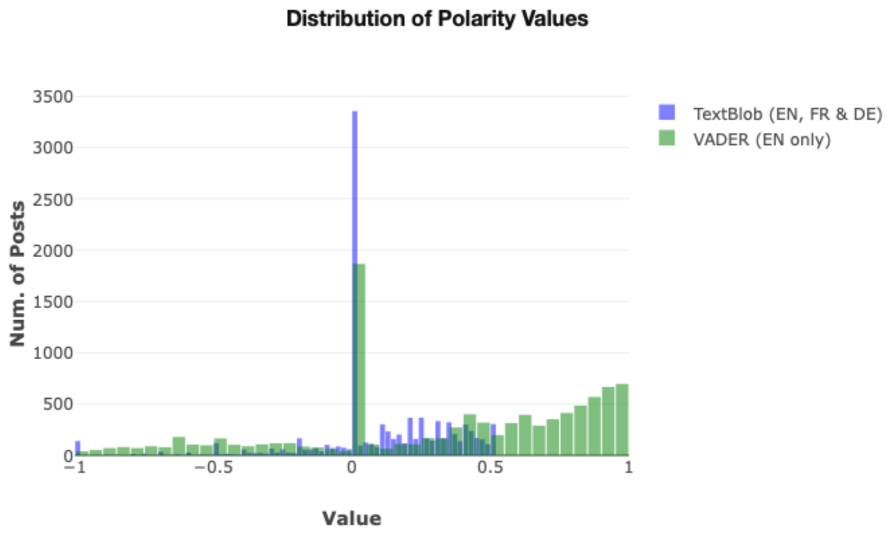

Figure 9. Distribution of Polarity Values

Figure 9 shows the distribution of polarity values. Based on the distribution of polarity values, it is evident that VADER and TextBlob agree on categorizing 9,579 (73.34%) out of 13,061 English language posts. This level of agreement is considered moderate, as indicated by a Cohen's kappa statistic of 0.464. Such a statistic reflects a fair level of consistency between the two models, though there remains room for discrepancies. These findings underscore the importance of utilizing multiple models to ensure robust sentiment analysis. Ultimately, this moderate agreement highlights the need for continuous refinement in sentiment classification methodologies to achieve greater accuracy and reliability.

Expressly, both VADER and TextBlob libraries agree on the classification of sentiment in several posts: 845 (8.82%) posts exhibit negative sentiments with polarity scores of -0.05 or lower, 1,349 (14.08%) posts are categorized as having neutral sentiments with polarity scores between -0.05 and 0.05, and 7,385 (77.10%) posts show positive

Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 411-420
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5416

sentiments with polarity scores of 0.05 or higher. This agreement underscores the robustness of these sentiment analysis models in identifying general sentiment trends across a substantial dataset. Such consistency between the models enhances confidence in the reliability of the sentiment classification process. Ultimately, this alignment supports using these tools in comprehensive sentiment analysis, contributing to more accurate and actionable insights.

The analysis reveals discrepancies between VADER and TextBlob in sentiment classification for a subset of posts. Specifically, there are 378 posts where VADER assigned positive polarity scores, while TextBlob assigned negative scores. Conversely, there are 396 posts where VADER assigned negative polarity scores and TextBlob assigned positive scores. These divergences highlight the inherent differences in the algorithms and approaches used by the two models. Such variations underscore the necessity of employing multiple sentiment analysis tools to capture a more comprehensive picture of user sentiment. Ultimately, understanding and addressing these discrepancies are crucial for enhancing the accuracy and reliability of sentiment analysis in digital content evaluation.

3.3 Discussion

The relationship between toxicity and sentiment analysis in tourism and travel content reviews, particularly video reviews, is pivotal in understanding user engagement and digital marketing effectiveness. Toxicity analysis helps identify negative interactions that could deter potential tourists, while sentiment analysis provides insights into the overall emotional response to the content. When applied to video reviews, these analytical tools reveal how users perceive and react to travel destinations, influencing engagement levels. Consequently, integrating toxicity and sentiment analysis into digital marketing strategies enables marketers to tailor content to foster positive interactions and mitigate negative feedback. Ultimately, this comprehensive approach enhances user engagement and optimizes the impact of digital marketing efforts in the tourism sector.

A comprehensive understanding of tourist perception and behavior is achieved through sentiment analysis. By categorizing and evaluating the emotional tone of user-generated content, insights into tourists' attitudes and preferences become evident. These insights reveal the overall satisfaction levels and specific aspects of the travel experience that resonate most with tourists [34]. Consequently, sentiment analysis is a valuable tool for identifying trends and patterns in tourist behavior [35]. Ultimately, leveraging these insights enables the development of targeted strategies to enhance tourist satisfaction and influence positive behavioral outcomes.

The narrative of travel content, when structured in a well-designed storyboard aimed at capturing viewers' attention and simplifying content comprehension, significantly enhances the intention to visit travel destinations. An engaging and coherent narrative not only captivates the audience but also effectively communicates the unique aspects and attractions of the destination [36]. The content positively influences viewers' perceptions by fostering a deeper connection and understanding [37]. Consequently, a meticulously crafted travel narrative is crucial in motivating potential tourists to visit showcased destinations [38]. Ultimately, this approach underscores the importance of strategic content creation in boosting tourism engagement and visit intentions.

The limitation of this research lies in its reliance on sentiment analysis models such as VADER and TextBlob, which may not fully capture the complexities of human emotions and cultural nuances in travel content reviews. Additionally, the analysis is constrained by the accuracy and comprehensiveness of the datasets used, which might not encompass the entire spectrum of user interactions and sentiments [39]. The potential biases inherent in these models and datasets could affect the reliability of the findings [40]. Therefore, future studies should consider incorporating more advanced and diverse analytical tools to address these limitations. Ultimately, acknowledging these constraints is essential for refining methodologies and enhancing the validity of sentiment analysis in tourism research.

Future research should explore integrating more sophisticated sentiment analysis models and diverse datasets to better capture the intricacies of user emotions and cultural contexts in travel content reviews. Expanding the scope of analysis to include multilingual datasets and utilizing advanced machine learning techniques could provide a more comprehensive understanding of global tourist perceptions. Additionally, investigating the impact of various content formats and platforms on user engagement and sentiment would offer valuable insights. Therefore, these recommendations aim to enhance the robustness and applicability of sentiment analysis in the tourism sector. Ultimately, such advancements will contribute to more effective and targeted digital marketing strategies.

4. CONCLUSION

In conclusion, the Tourism and Travel Content Analysis (TTCA) framework has proven instrumental in providing a comprehensive understanding of user sentiment and behavior in response to digital travel content. By employing sentiment analysis models such as VADER and TextBlob, this research has highlighted key trends in tourist perceptions and interactions with travel-related videos. The sentiment classification results, based on the analysis of 13,162 posts using VADER, showed 1,832 (13.92%) negative, 1,977 (15.02%) neutral, and 9,353 (71.06%) positive sentiments. Similarly, TextBlob analysis revealed 1,378 (10.47%) negative, 3,489 (26.51%) neutral, and 8,295 (63.02%) positive sentiments for English posts. The toxicity scores calculated using the Detoxify and Perspective models indicated a range from low to high levels of toxic content, with specific metrics such as 0.09913 to 0.96069 for toxicity, 0.00844 to 0.61261 for severe toxicity, and various other scores for identity attack, insult, profanity, and threat. The findings underscore the importance of well-crafted narratives in influencing tourist behavior and visit intentions while highlighting the limitations of the current models in capturing emotional and cultural nuances. Future research should incorporate more advanced analytical tools and diverse datasets to address these limitations. Ultimately, the TTCA framework offers valuable insights that enhance digital marketing strategies and improve engagement within the tourism sector.

ACKNOWLEDGMENT

Thanks to the Tourism Study Program, Faculty of Business Administration and Communication, Atma Jaya Catholic University of Indonesia, LPPM, and Center of Digital Transformation and Tourism Development for the support and invaluable contributions to this research publication.

REFERENCES

,[1] J. Sihvonen and L. L. M. Turunen, "Multisensory experiences at travel fairs: What evokes feelings of pleasure, arousal and dominance among visitors?," *J. Conv. Event Tour.*, vol. 23, no. 1, pp. 63–85, 2022, doi: 10.1080/15470148.2021.1949417.

[2] C. Yanhao, H. Mustafa, and T. P. Ling, "Mapping the research of travel live streaming: A bibliometric analysis using Vosviewer," *Cogent Soc. Sci.*, vol. 10, no. 1, 2024, doi: 10.1080/23311886.2023.2285254.

[3] S. W. Lin, K. F. Wang, and Y. H. Chiu, "Effects of tourists' psychological perceptions and travel choice behaviors on the nonmarket value of urban ecotourism during the COVID-19 pandemic- case study of the Maokong region in Taiwan," *Cogent Soc. Sci.*, vol. 8, no. 1, 2022, doi: 10.1080/23311886.2022.2095109.

[4] C. (Cherise) Li, E. Agyeiwaah, A. Imboden, and Y. M. Lee, "Understanding young Chinese customers' perceptions of unsafe city image restoration after Covid-19 outbreak," *Int. J. Tour. Cities*, vol. 9, no. 4, pp. 940–959, Jan. 2023, doi: 10.1108/IJTC-12-2022-0265.

[5] N. Mohammad, M. H. Hanafiah, and M. S. M. Zahari, "The influence of heritage food consumption and experiential value on Terengganu's destination image and tourists' behavioural intention," *J. Herit. Tour.*, vol. 17, no. 6, pp. 685–701, 2022, doi: 10.1080/1743873X.2022.2110886.

[6] S. M. Fazal-e-Hasan, G. Mortimer, H. Ahmadi, M. Abid, O. Farooque, and A. Amrollahi, "How tourists' negative and positive emotions motivate their intentions to reduce food waste," *J. Sustain. Tour.*, vol. 0, no. 0, pp. 1–21, 2023, doi: 10.1080/09669582.2023.2264539.

[7] X. Chen and Z. Chen, "Impact of video content on gastronomic image construction and tourists' intention to (re-) visit Macao," *Tour. Recreat. Res.*, vol. 0, no. 0, pp. 1–15, 2023, doi: 10.1080/02508281.2023.2296812.

[8] A. Ben Aicha and R. Bouzaabia, "The effects of video storytelling advertising on consumers' online reactions on Facebook: a cross cultural study," *Qual. Mark. Res.*, vol. 26, no. 3, pp. 247–268, Jan. 2023, doi: 10.1108/QMR-06-2022-0101.

[9] D. Nuranindya, A. R. Berto, and A. Y. M. Wahyu, "Optimisme Saat Pandemi : Strategi Storytelling Pada Video Promosi Destinasi Pariwisata Super Prioritas," *J. Stud. Komun.*, vol. 7, no. 2, pp. 637–661, 2023, doi: 10.25139/jsk.v7i2.

[10] S. Lee, Y. Chang, O. K. D. Lee, S. Ryu, and Q. Yin, "Exploring online social platform affordances for digital creators: a multi-method approach using qualitative and configurational analysis," *Ind. Manag. Data Syst.*, vol. 124, no. 4, pp. 1501–1530, Jan. 2024, doi: 10.1108/IMDS-12-2023-0951.

[11] S. P. Hammond, N. J. Cooper, and P. Jordan, "Mental health, identity and informal education opportunities for adolescents with experience of living in state care: a role for digital storytelling," *Cambridge J. Educ.*, vol. 51, no. 6, pp. 713–732, 2021, doi: 10.1080/0305764X.2021.1919057.

[12] J. Dennis and S. Sampaio-Dias, "'Tell the Story as You'd Tell It to Your Friends in a Pub': Emotional Storytelling in Election Reporting by BuzzFeed News and Vice News," *Journal. Stud.*, vol. 22, no. 12, pp. 1608–1626, 2021, doi: 10.1080/1461670X.2021.1910541.

[13] H. Devan *et al.*, "'Power of Storytelling': A Content Analysis of Chronic Pain Narratives on YouTube," *Can. J. Pain*, vol. 5, no. 1, pp. 117–129, 2021, doi: 10.1080/24740527.2021.1929117.

[14] D. Paiva, L. Carvalho, E. Brito-Henriques, A. M. Sousa, A. L. Soares, and S. T. Azambuja, "Digital storytelling and hopeful last chance tourism experiences," *Tour. Geogr.*, vol. 25, no. 5, pp. 1428–1444, 2023, doi: 10.1080/14616688.2023.2224043.

[15] P. Tosun, A. Uslu, and E. Erul, "Connecting through chatbots: residents' insights on digital storytelling, place attachment, and value co-creation," *Curr. Issues Tour.*, pp. 1–24, 2024, doi: 10.1080/13683500.2024.2316857.

[16] N. Nicoli, K. Henriksen, M. Komodromos, and D. Tsagalas, "Investigating digital storytelling for the creation of positively engaging digital content," *EuroMed J. Bus.*, vol. 17, no. 2, pp. 157–173, Jan. 2022, doi: 10.1108/EMJB-03-2021-0036.

[17] T. Leinonen, J. Brinck, H. Vartiainen, and N. Sawhney, "Augmented reality sandboxes: children's play and storytelling with mirror worlds," *Digit. Creat.*, vol. 32, no. 1, pp. 38–55, 2021, doi: 10.1080/14626268.2020.1868535.

[18] A. Terah Ambala, "Adapting Digital Storytelling Methods to African Contexts: Insights from the Utaifa Project," *African Journal. Stud.*, vol. 41, no. 4, pp. 22–34, 2020, doi: 10.1080/23743670.2020.1840413.

[19] M. Zayani, "Digital Journalism, Social Media Platforms, and Audience Engagement: The Case of AJ+," *Digit. Journal.*, vol. 9, no. 1, pp. 24–41, 2021, doi: 10.1080/21670811.2020.1816140.

[20] K. K. Coker, R. L. Flight, and D. M. Baima, "Video storytelling ads vs argumentative ads: how hooking viewers enhances consumer engagement," *J. Res. Interact. Mark.*, vol. 15, no. 4, pp. 607–622, Jan. 2021, doi: 10.1108/JRIM-05-2020-0115.

[21] N. Churchill, "Development of students' digital literacy skills through digital storytelling with mobile devices," *EMI. Educ. Media Int.*, vol. 57, no. 3, pp. 271–284, 2020, doi: 10.1080/09523987.2020.1833680.

[22] E. E. Vazquez, "Effects of enduring involvement and perceived content vividness on digital engagement," *J. Res. Interact. Mark.*, vol. 14, no. 1, pp. 1–16, Jan. 2020, doi: 10.1108/JRIM-05-2018-0071.

[23] R. Sann and P.-C. Lai, "Do expectations towards Thai hospitality differ? The views of English vs Chinese speaking travelers," *Int. J. Cult. Tour. Hosp. Res.*, vol. 15, no. 1, pp. 43–58, Jan. 2021, doi: 10.1108/IJCTHR-01-2020-0010.

Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 411-420
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5416

[24] Y. S. Yang, M. J. Seo, and J. Byun, "The Experience of Video Interpretation and Satisfaction," *Tour. Plan. Dev.*, pp. 1–15, 2024, doi: 10.1080/21568316.2024.2306350.
[25] N. Arora and S. Lata, "YouTube channels influence on destination visit intentions," *J. Indian Bus. Res.*, vol. 12, no. 1, pp. 23–42, Jan. 2020, doi: 10.1108/JIBR-09-2019-0269.
[26] A. Zaib Abbasi *et al.*, "Tourism promotion through vlog advertising and customer engagement behaviours of generation Z," *Curr. Issues Tour.*, vol. 26, no. 22, pp. 3651–3670, 2023, doi: 10.1080/13683500.2022.2144156.
[27] C. M. Segijn, E. Maslowska, T. Araujo, and V. Viswanathan, "Engaging with TV events on Twitter: The interrelations between TV consumption, engagement actors, and engagement content," *Internet Res.*, vol. 30, no. 2, pp. 381–401, Jan. 2020, doi: 10.1108/INTR-08-2018-0389.
[28] W. Manyanga, J. Kanyepe, L. Chikazhe, and T. Manyanga, "The effect of social media marketing on brand loyalty in the hospitality industry in Zimbabwe: the moderating role of age," *Cogent Bus. Manag.*, vol. 11, no. 1, p., 2024, doi: 10.1080/23311975.2024.2302311.
[29] L. W. Liu, C. C. Wang, P. Pahrudin, A. F. Royanow, C. Lu, and I. Rahadi, "Does virtual tourism influence tourist visit intention on actual attraction? A study from tourist behavior in Indonesia," *Cogent Soc. Sci.*, vol. 9, no. 1, 2023, doi: 10.1080/23311886.2023.2240052.
[30] P. M. Lau, J. S. Y. Ho, and P. Pillai, "Research note – Sensational museums on TikTok: reaching young virtual tourists with short videos," *Consum. Behav. Tour. Hosp.*, vol. 19, no. 1, pp. 70–81, Jan. 2024, doi: 10.1108/CBTH-04-2023-0039.
[31] M. Li, M. Cheng, V. Quintal, and I. Cheah, "From live streamer to viewer : exploring travel live streamer persuasive linguistic styles and their impacts on travel intentions," *J. Travel Tour. Mark.*, vol. 40, no. 8, pp. 764–777, 2023, doi: 10.1080/10548408.2023.2294071.
[32] X. Ma, J. Zhang, Y. Sun, P. Wang, and R. Tang, "Awakening travel intentions by food travel vlogs: far from the sky and close to the eyes," *Leis. Stud.*, vol. 00, no. 00, pp. 1–16, 2023, doi: 10.1080/02614367.2023.2293023.
[33] J. H. Kim, F. Hu, J. Wen, and H. Hou, "Understanding tourists' dining behaviors at traditional Chinese nutraceutical restaurants," *Anatolia*, vol. 00, no. 00, pp. 1–25, 2023, doi: 10.1080/13032917.2023.2282041.
[34] Z. Gao *et al.*, "Exploring tourist spatiotemporal behavior differences and tourism infrastructure supply–demand pattern fusing social media and nighttime light remote sensing data," *Int. J. Digit. Earth*, vol. 17, no. 1, pp. 1–22, 2024, doi: 10.1080/17538947.2024.2310723.
[35] H. Han, "Consumer behavior and environmental sustainability in tourism and hospitality: a review of theories, concepts, and latest research," *J. Sustain. Tour.*, vol. 29, no. 7, pp. 1021–1042, 2021, doi: 10.1080/09669582.2021.1903019.
[36] Y. Gao, Y. Ma, K. Bai, Y. Li, and X. Liu, "Which factors influence individual pro-environmental behavior in the tourism context: rationality, affect, or morality?," *Asia Pacific J. Tour. Res.*, vol. 26, no. 5, pp. 516–538, 2021, doi: 10.1080/10941665.2021.1876117.
[37] L. Perpiña, L. Prats, and R. Camprubí, "Image and risk perceptions: an integrated approach," *Curr. Issues Tour.*, vol. 24, no. 3, pp. 367–384, 2021, doi: 10.1080/13683500.2020.1715355.
[38] K. Thirumaran, Z. Mohammadi, S. Azzali, E. L. Eijdenberg, and G. Donough-Tan, "Transformed landscapes, tourist sentiments: the place making narrative of a luxury heritage hotel in Singapore," *J. Herit. Tour.*, vol. 18, no. 2, pp. 243–264, 2023, doi: 10.1080/1743873X.2023.2169151.
[39] G. Rasool and A. Pathania, "Reading between the lines: untwining online user-generated content using sentiment analysis," *J. Res. Interact. Mark.*, vol. 15, no. 3, pp. 401–418, Jan. 2021, doi: 10.1108/JRIM-03-2020-0045.
[40] F. Mehraliyev, I. C. C. Chan, and A. P. Kirilenko, "Sentiment analysis in hospitality and tourism: a thematic and methodological review," *Int. J. Contemp. Hosp. Manag.*, vol. 34, no. 1, pp. 46–77, 2022, doi: 10.1108/IJCHM-02-2021-0132.