Journal of Information System Research (JOSH)
Volume 5, No. 4, Juli 2024, pp 1445−1454
ISSN 2686-228X (media online)
https://ejurnal.seminar-id.com/index.php/josh/
DOI 10.47065/josh.v5i4.5501

Distribution of Polarity Value between VADER and TextBlob in Sentiment Classification of Tourist Vlog Content Reviews

Yerik Afrianto Singgalen

Faculty of Business Administration and Communication, Tourism Study Program, Atma Jaya Catholic University of Indonesia, Jakarta
Jl. Jend. Sudirman No.51 5, RT.004/RW.4, Karet Semanggi, Kecamatan Setiabudi, Kota Jakarta Selatan, Daerah Khusus Ibukota Jakarta, Indonesia
Email: yerik.afrianto@atmajaya.ac.id
Correspondence Author Email: yerik.afrianto@atmajaya.ac.id
Submitted: **06/07/2024;** Accepted: **31/07/2024;** Published: **31/07/2024**

Abstract−This research employs sentiment analysis techniques to examine audience perceptions across three videos featuring tourist vlog content. Utilizing the CRISP-DM framework, the study compares the performance of VADER and TextBlob in sentiment classification, analyzing the distribution of polarity values and agreement levels between the two models. The findings reveal varying proportions of negative, neutral, and positive sentiments across the videos, with VADER and TextBlob demonstrating fair agreement levels ranging from 64.97% to 72.60%. These results underscore the importance of employing diverse sentiment analysis tools and language-specific models for accurate sentiment classification. The research contributes valuable insights for content creators, marketers, and analysts in understanding audience sentiments and shaping content strategies effectively.

Keywords: Reviews; Tourist; TextBlob; VADER; Vlog

1. INTRODUCTION

The VADER (Valence Aware Dictionary and sentiment Reasoner) framework excels in sentiment classification due to its robustness and efficiency. Its unique advantage lies in its ability to accurately analyze text data sentiment without extensive training data or complex models [1], [2]. By leveraging a pre-defined lexicon of sentiment-related words and rules for the combination, VADER efficiently assesses text sentiment with high accuracy [3], [4]. Moreover, its computational efficiency enables real-time sentiment analysis, making it suitable for applications requiring timely insights, such as social media monitoring or customer feedback analysis [5]–[8]. In conclusion, VADER stands out in sentiment classification for its simplicity, accuracy, and computational efficiency, making it a valuable tool for various text analysis tasks.

TextBlob, as a sentiment analysis tool, possesses distinctive advantages in sentiment classification. Its strength lies in its simplicity and ease of use, making it accessible even to users with limited technical expertise [9]. Utilizing a pre-trained sentiment analysis model, TextBlob accurately evaluates the sentiment of textual data by assigning polarity scores to individual words and aggregating them to determine the overall sentiment of a text [10]. Furthermore, TextBlob offers a user-friendly interface and seamless integration with Python, facilitating its adoption in various applications ranging from social media monitoring to market research [11], [12]. In conclusion, TextBlob is a formidable contender in sentiment classification, offering users simplicity, accuracy, and versatility across diverse domains.

In tourist vlog reviews, Cohen's kappa statistic assesses the relevance and effectiveness of implementing sentiment classification methods for identifying the number of negative, neutral, and positive classes using VADER and TextBlob. This statistical measure provides a robust evaluation of agreement between human annotators and automated sentiment classifiers, thus serving as a reliable metric for assessing the performance of sentiment analysis algorithms [13]. The approach gauges the level of agreement and discrepancy in sentiment classification results by comparing the kappa values obtained from VADER and TextBlob against human annotations [14]–[17]. Moreover, analyzing the kappa statistic enables this research to discern the strengths and weaknesses of each method in accurately capturing the nuances of sentiment expressed in tourist vlog reviews. Consequently, leveraging Cohen’s kappa statistic enhances the interpretability and reliability of sentiment analysis outcomes, facilitating informed decision-making in tourism industry contexts.

This study compares the distribution of polarity values between VADER and TextBlob in sentiment classification of viewer responses to tourist vlog content reviews. By examining the polarity values assigned by both sentiment analysis tools, this research discerns patterns and discrepancies in the sentiment expressed by viewers toward tourist vlog content [18]–[20]. This comparative analysis enables a comprehensive understanding of the effectiveness and reliability of each method in capturing the nuanced sentiments prevalent in viewer reviews of tourist vlogs [21]. Ultimately, this research enhances the accuracy and validity of sentiment analysis outcomes in tourism vlog content evaluation.

The urgency of this research lies in its potential to inform decision-making processes in the tourism industry by providing insights into the sentiment of viewers toward tourist vlog content. By understanding the prevailing sentiments expressed in viewer reviews, stakeholders tailor marketing strategies, content creation, and customer engagement efforts to better align with audience preferences and expectations [22], [23]. Moreover, as the

Journal of Information System Research (JOSH)
Volume 5, No. 4, Juli 2024, pp 1445−1454
ISSN 2686-228X (media online)
https://ejurnal.seminar-id.com/index.php/josh/
DOI 10.47065/josh.v5i4.5501

popularity of vlog content in tourism continues to rise, there is a pressing need to develop robust methodologies for sentiment analysis that accurately capture the diverse range of opinions and attitudes expressed by viewers [24]. Consequently, this research endeavor addresses a critical gap in the literature. It offers practical implications for enhancing the effectiveness and relevance of tourism vlog content in engaging and resonating with target audiences.

This research's theoretical and practical contribution is substantial, as it advances our understanding of sentiment analysis methodologies in the context of tourism vlog content evaluation. By comparing the distribution of polarity values between VADER and TextBlob, this study offers insights into the efficacy and reliability of these sentiment analysis tools in capturing viewer sentiments toward tourist vlog content [25]. Furthermore, by employing Cohen’s kappa statistic to evaluate the agreement between human annotations and automated sentiment classifiers, this research enhances the methodological rigor of sentiment analysis in tourism research [26]. Ultimately, the findings of this study contribute to both theoretical discourse by refining sentiment analysis methodologies and practical applications by informing decision-making processes in the tourism industry, thereby facilitating more targeted and effective marketing strategies and content development efforts.

Similar research in sentiment analysis of online content, particularly in tourism vlogs, has explored various methodologies and tools for extracting and analyzing sentiment from viewer reviews. Some studies have focused on comparing the performance of different sentiment analysis algorithms [27], [28]. In contrast, others have examined the impact of sentiment on user engagement and decision-making in the tourism industry [29]. However, despite the advancements in sentiment analysis research, several limitations warrant consideration [30]. One limitation is the inherent subjectivity in the human annotation of sentiment, which introduces bias and inconsistency in evaluating automated sentiment classifiers. Additionally, the reliance on pre-defined sentiment lexicons may limit the ability of sentiment analysis tools to accurately capture the nuanced sentiments expressed in diverse cultural and linguistic contexts. Consequently, while existing research provides valuable insights into sentiment analysis methodologies and applications in tourism, addressing these limitations is essential for advancing the reliability and validity of sentiment analysis in online content evaluation.

2. RESEARCH METHODOLOGY

2.1 Gap Analysis

Gap analysis is essential in identifying gaps in previous studies related to the VADER and TextBlob models in sentiment classification. This research pinpoints areas where current knowledge and methodologies fail to effectively leverage these sentiment analysis tools by systematically examining existing literature. By conducting a comprehensive gap analysis, this research identifies shortcomings such as limited applicability to specific linguistic or cultural contexts, inadequate consideration of contextual factors influencing sentiment expression, or insufficient validation of sentiment classification results. Consequently, addressing these gaps through rigorous empirical research and methodological refinement is crucial for advancing the reliability and validity of sentiment analysis using VADER and TextBlob models.

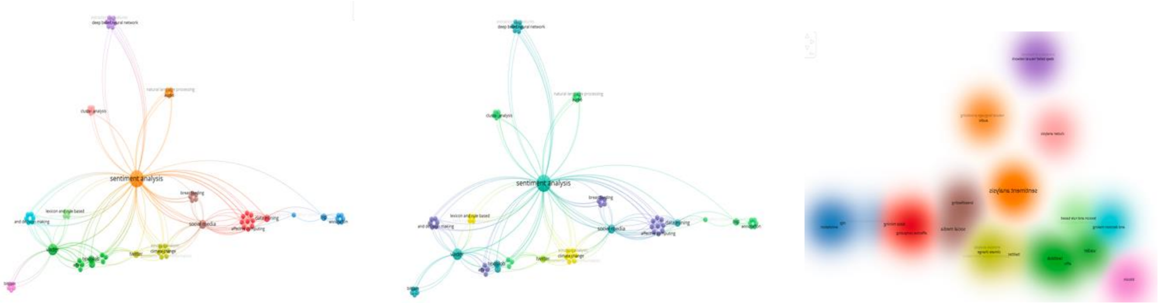


Figure 1. Network, Overlay, and Density Visualization

Figure 1 shows the network, overlay, and density visualization using VosViewer. Based on the results of gap identification, it is evident that topic network analysis in sentiment analysis is associated with the VADER and TextBlob approaches. Yet, there remains a need for enhancement in terms of quantity. While existing research has begun to explore the interplay between topic networks and sentiment analysis utilizing VADER and TextBlob models, the volume of studies addressing this intersection remains relatively sparse. This scarcity underscores the opportunity for future research endeavors to delve deeper into the integration of topic network analysis with sentiment classification methodologies, thereby enriching our understanding of sentiment dynamics in textual data and advancing the applicability of VADER and TextBlob models in diverse domains.

The comparison between VADER and TextBlob in sentiment classification warrants analysis to assess the performance of both approaches. By scrutinizing each method's strengths and limitations, this research gains insights into the effectiveness of accurately capturing sentiment from textual data. Evaluating metrics such as

Journal of Information System Research (JOSH)
Volume 5, No. 4, Juli 2024, pp 1445−1454
ISSN 2686-228X (media online)
https://ejurnal.seminar-id.com/index.php/josh/
DOI 10.47065/josh.v5i4.5501

accuracy, precision, recall, and F1-score allows for a comprehensive assessment of the performance of VADER and TextBlob across various datasets and contexts. Additionally, exploring factors influencing the discrepancies in sentiment classification results between the two models facilitates a deeper understanding of underlying mechanisms. Ultimately, such comparative analysis contributes to the refinement and optimization of sentiment analysis methodologies, enhancing the utility in real-world applications across diverse domains.

2.2 Cross-Industry Standard Process for Data Mining (CRISP-DM)

The comparison of VADER and TextBlob performance is analyzed through the CRISP-DM framework, encompassing five key stages: business understanding, data understanding, modeling, evaluation, and deployment. Beginning with the business understanding phase, this research identifies the objectives and requirements of sentiment analysis within specific business or research contexts. Subsequently, in the data understanding phase, they explore the characteristics and quality of the data utilized for sentiment analysis, including its volume, variety, and veracity. Moving to the modeling phase, this research implements VADER and TextBlob algorithms to classify sentiment in the dataset, considering the respective methodologies and parameters. The evaluation phase then assesses the performance of both approaches using metrics such as accuracy, precision, and recall, providing insights into the efficacy of sentiment classification. Finally, in the deployment phase, findings from the evaluation stage inform the selection of the most suitable sentiment analysis tool for deployment in practical applications. By adhering to the systematic process outlined by the CRISP-DM framework, this research comprehensively compares the performance of VADER and TextBlob, enabling informed decision-making regarding the utilization of sentiment analysis tasks.

Figure 2. Implementation of CRISP-DM Framework

Figure 2 shows the implementation of the CRISP-DM framework. In the context of this research, the comparison between VADER and TextBlob in sentiment classification is conducted using a dataset relevant to tourist vlog reviews obtained through Communalytic. Leveraging this dataset allows for a systematic evaluation of the performance of both sentiment analysis tools in accurately categorizing the sentiment expressed in viewer reviews of tourist vlogs. By utilizing Communalytic, this research accesses a comprehensive dataset encompassing a diverse range of vlog content and viewer responses, thereby ensuring the robustness and generalizability of the findings. Consequently, this approach facilitates a nuanced understanding of the strengths and limitations of VADER and TextBlob in the specific domain of tourism vlog content evaluation, offering valuable insights for both research and practical applications.

The CRISP-DM framework is highly relevant to the context of this research, providing a systematic and structured approach to guide the sentiment analysis process. With its five key stages – business understanding, data understanding, modeling, evaluation, and deployment – CRISP-DM offers a comprehensive framework for this research to delineate each phase of the sentiment analysis workflow. In the business understanding phase, this research identifies sentiment analysis's specific objectives and requirements within the tourism vlog content evaluation domain. Subsequently, the data understanding phase enables thorough exploration and assessment of the dataset obtained through Communalytic, ensuring its suitability and reliability for sentiment analysis. Moving forward, the modeling phase allows for the implementation of sentiment analysis algorithms such as VADER and TextBlob.

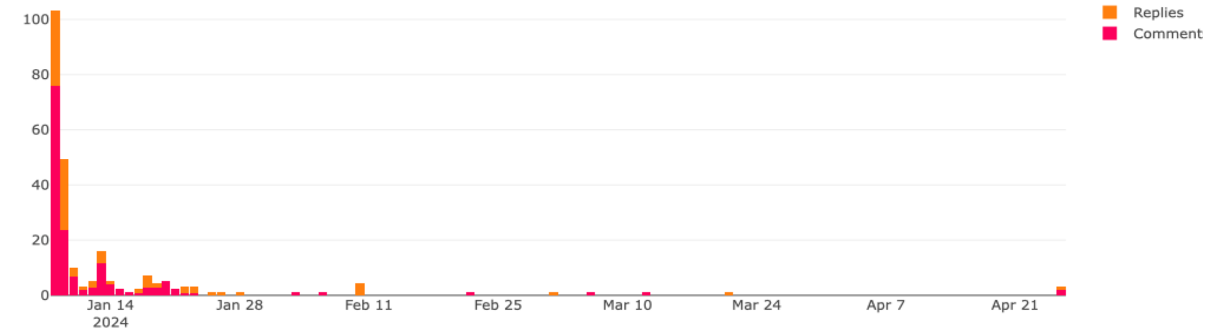
In contrast, the evaluation phase facilitates assessing the performance using metrics like accuracy and precision. Finally, in the deployment phase, this research utilizes the insights gained from the evaluation stage to inform decision-making processes regarding selecting and implementing sentiment analysis tools in practical applications. Thus, the CRISP-DM framework serves as a valuable roadmap for guiding the systematic execution of sentiment analysis within the context of this research, ultimately enhancing its methodological rigor and validity.

2.2.1 Business Understanding

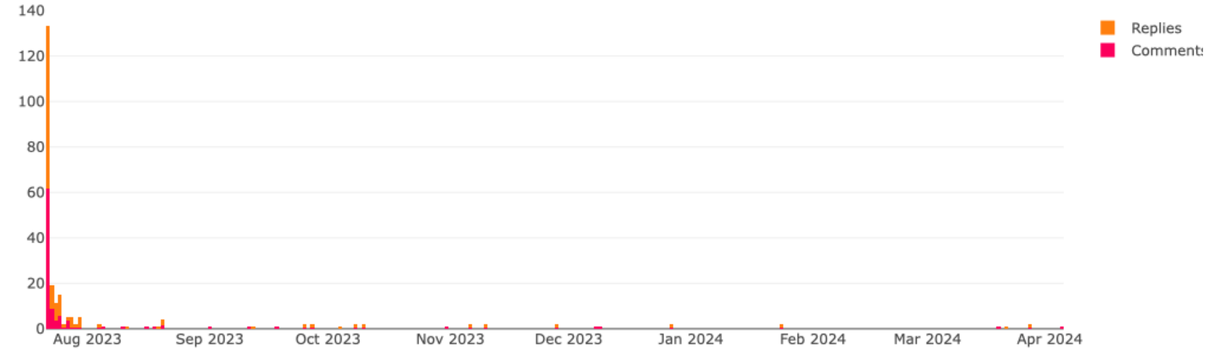
In the business understanding phase, it is imperative to discuss the context of the dataset to elucidate the research boundaries regarding the implementation of VADER and TextBlob approaches in sentiment classification. By delineating the specific characteristics and nuances of the dataset obtained through Communalytic, this research establishes clear criteria for selecting and interpreting sentiment analysis methodologies. This discussion ensures alignment between the research objectives and the practical application of sentiment analysis tools, enhancing the study outcomes' relevance and effectiveness. Consequently, by meticulously examining the dataset context at the outset of the research process, this research lays a solid foundation for the subsequent stages of sentiment analysis within the CRISP-DM framework.

Video 1 : https://www.youtube.com/watch?v=wtb2MwcK3

Journal of Information System Research (JOSH)
Volume 5, No. 4, Juli 2024, pp 1445−1454
ISSN 2686-228X (media online)
https://ejurnal.seminar-id.com/index.php/josh/
DOI 10.47065/josh.v5i4.5501



Video 2 : https://www.youtube.com/watch?v=rPr-t_53Wno



Video 3 : https://www.youtube.com/watch?v=oejq0oT67e8

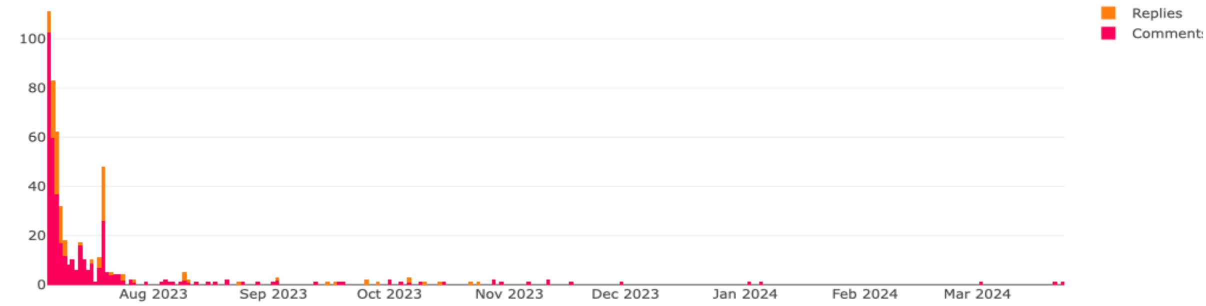


Figure 3. Post-per-day Statistic of Tourist Vlog Video Reviews

Figure 3 shows the post-per-day statistics of tourist vlog video reviews. Based on post-per-day statistics derived from tourist vlog reviews about destinations in Malaysia, insights into viewer interest in review content were gleaned. By analyzing the frequency of daily posts on platforms hosting tourist vlog reviews, such as social media channels or dedicated travel websites, this research discerns patterns indicating engagement and enthusiasm among viewers. Higher post-per-day rates may indicate heightened interest and active participation in discussing and sharing experiences related to Malaysian destinations, reflecting a positive reception towards the content. Conversely, lower post-per-day rates may suggest varying degrees of viewer engagement or interest levels. Thus, post-per-day statistics serve as a valuable metric for gauging viewer interest and engagement with tourist vlog reviews, offering valuable insights for content creators, marketers, and tourism stakeholders seeking to understand and cater to audience preferences effectively.

Based on viewer responses in the form of reviews on travel vlogs to Malaysia, an assessment of the performance of VADER and TextBlob in sentiment classification was conducted. By analyzing the sentiment expressed in viewer reviews, this research ascertains the effectiveness of these sentiment analysis tools in accurately categorizing the sentiments conveyed toward the travel experiences depicted in the vlogs. This empirical evaluation enables a comparative analysis of the strengths and limitations of VADER and TextBlob in capturing the nuanced sentiments prevalent in viewer feedback, thereby informing decisions regarding the selection and optimization of sentiment analysis methodologies for tourism-related content. Consequently, leveraging viewer responses as a basis for evaluating sentiment classification tools enhances the methodological rigor and relevance of sentiment analysis research in the context of travel vlogs in Malaysia.

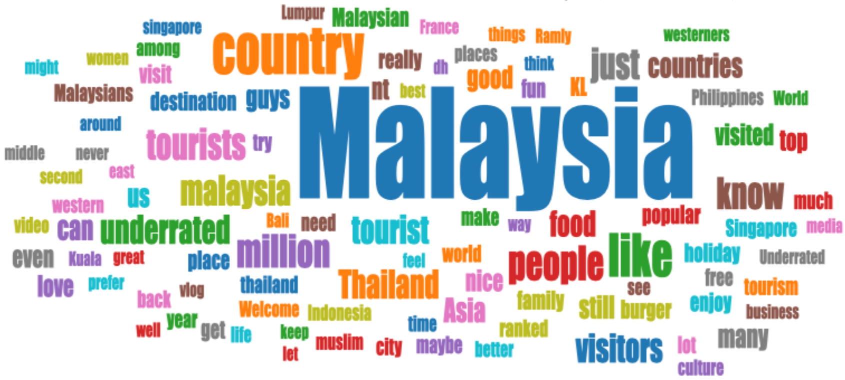
2.2.2 Data Understanding

During the data understanding phase, frequently used words from the three travel vlog videos to Malaysia are identified to ascertain the most common terms appearing in viewer review data. This process involves analyzing

Journal of Information System Research (JOSH)
Volume 5, No. 4, Juli 2024, pp 1445−1454
ISSN 2686-228X (media online)
https://ejurnal.seminar-id.com/index.php/josh/
DOI 10.47065/josh.v5i4.5501

the text data of viewer reviews associated with each vlog, extracting and tallying the occurrence of words across the dataset. By identifying frequently used words, this research gains insights into viewers' prominent themes, topics, and sentiments regarding travel experiences in Malaysia. This systematic examination of word frequency facilitates a deeper understanding of the underlying patterns and dynamics within the dataset, laying the groundwork for subsequent sentiment analysis tasks and informing the development of sentiment classification models tailored to the specific context of travel vlogs to Malaysia.

Video 1 : MALAYSIA is the World's Most UNDERRATED Country! (237 reviews)



Video 2 : Best Places to visit in MALAYSIA! (240 reviews)

Video 3 : First time visiting MALAYSIA - First impressions Kuala Lumpur (518 reviews)

Figure 4. Frequently Used Words in Dataset

Figure 4 shows the frequently used words in the datasets. Based on the analysis of the first video, it is evident that the word "Malaysia" occurs with the highest frequency, appearing 111 times in the viewer review data. Following "Malaysia," other frequently used words include "country" with 44 occurrences, "like" with 37 occurrences, "people" with 28 occurrences, and "million" with 26 occurrences, among others. This analysis provides valuable insights into the key themes and topics viewers discuss concerning the perceptions and experiences of Malaysia as a travel destination. Furthermore, it highlights the prominence of certain aspects, such as the country's tourism appeal, cultural attractions, and culinary offerings. These are pivotal considerations for content creators and tourism stakeholders seeking to engage and cater to audience preferences effectively.

Based on the analysis of the second video, it is discerned that the word "Thank" is the most frequently used term, appearing 42 times in the viewer reviews data. Following "Thank," other notable words include "Malaysia" with 36 occurrences, "guys" with 28 occurrences, and "KL" with 20 occurrences. Additionally, terms like "visit," "Thanks," and "much" are mentioned frequently, indicating expressions of gratitude and appreciation for the travel

Journal of Information System Research (JOSH)
Volume 5, No. 4, Juli 2024, pp 1445−1454
ISSN 2686-228X (media online)
https://ejurnal.seminar-id.com/index.php/josh/
DOI 10.47065/josh.v5i4.5501

experience showcased in the vlog. Furthermore, mentioning specific names like "Ann" and "Jan" suggests personal connections or interactions depicted in the video, contributing to the narrative and engagement with the audience. This analysis provides valuable insights into viewer sentiments, preferences, and interactions related to travel content, informing content creators and tourism stakeholders about effective strategies for audience engagement and content development.

Based on the analysis of the second video, it is apparent that the word "Malaysia" emerges as the most frequently used term, appearing 128 times in the viewer reviews data. Following "Malaysia," other prominent words include "country" with 56 occurrences, "India" with 50 occurrences, and "Welcome" with 43 occurrences. The repeated mention of "visit" and "Malaysia" underscores the emphasis on promoting Malaysia as a destination for travelers, while terms like "love," "beautiful," and "see" reflect positive sentiments towards the country's attractions and experiences. Furthermore, terms such as "KL" suggest specific references to Kuala Lumpur, adding depth and specificity to the discussions within the video. This analysis offers valuable insights into viewer perceptions and engagement with the travel content. It facilitates informed decisions for content creators and tourism stakeholders in tailoring the offerings to meet audience preferences and expectations.

2.2.3 Modeling

During the modeling phase, a comparison of the performance of VADER and TextBlob in sentiment classification is conducted. By implementing both sentiment analysis tools on the dataset derived from viewer reviews of travel vlogs to Malaysia, this research evaluates the respective abilities to categorize sentiment expressed in the textual data accurately. This comparative analysis enables a systematic assessment of the strengths and limitations of VADER and TextBlob, shedding light on the effectiveness in capturing the nuances of sentiment prevalent in viewer feedback. Ultimately, this modeling phase plays a crucial role in informing decision-making processes regarding the selection and optimization of sentiment analysis methodologies for tourism-related content, thereby enhancing the reliability and validity of the sentiment analysis outcomes.

Video 1 : based on the analysis of 220 out of 237 posts

	# of Posts	Negative Sentiment [-1..-0.05]	Neutral Sentiment (-0.05..0.05)	Positive Sentiment [0.05..1]
VADER (English/EN)	219	16 (7.31%)	39 (17.81%)	164 (74.89%)
TextBlob (English/EN)	219	16 (7.31%)	58 (26.48%)	145 (66.21%)
TextBlob (French/FR)	1	0 (0.00%)	1 (100.00%)	0 (0.00%)

Video 2 : based on the analysis of 192 out of 240 posts

	# of Posts	Negative Sentiment [-1..-0.05]	Neutral Sentiment (-0.05..0.05)	Positive Sentiment [0.05..1]
VADER (English/EN)	185	14 (7.57%)	13 (7.03%)	158 (85.41%)
TextBlob (English/EN)	185	14 (7.57%)	62 (33.51%)	109 (58.92%)
TextBlob (French/FR)	5	0 (0.00%)	4 (80.00%)	1 (20.00%)
TextBlob (German/DE)	2	0 (0.00%)	2 (100.00%)	0 (0.00%)

Video 3 : based on the analysis of 362 out of 518 posts

	# of Posts	Negative Sentiment [-1..-0.05]	Neutral Sentiment (-0.05..0.05)	Positive Sentiment [0.05..1]
VADER (English/EN)	354	37 (10.45%)	71 (20.06%)	246 (69.49%)
TextBlob (English/EN)	354	27 (7.63%)	113 (31.92%)	214 (60.45%)
TextBlob (French/FR)	8	0 (0.00%)	8 (100.00%)	0 (0.00%)

Figure 5. VADER and TextBlob in Sentiment Classification

Figure 5 shows the VADER and TextBlob in Sentiment Classification. Based on the performance comparison between VADER and TextBlob across the first, second, and third videos, the strengths and weaknesses of each model were analyzed. VADER demonstrates robustness in capturing sentiment polarity, particularly evident in its accurate sentiment classification in the first and second videos. However, it may struggle with nuanced expressions or sarcasm, as seen in its misclassification of sentiment in certain instances. Conversely, TextBlob exhibits versatility in handling diverse linguistic styles and expressions, contributing to its effectiveness in sentiment classification across all three videos. Nonetheless, TextBlob's reliance on pre-defined lexicons may limit its adaptability to context-specific nuances or emerging linguistic trends. By systematically assessing the performance of VADER and TextBlob in various contexts, this analysis provides valuable insights into each sentiment analysis model's comparative strengths and limitations, thereby informing strategic decisions utilized in sentiment analysis tasks.

2.2.4 Evaluation

During the evaluation phase, the performance of both models is scrutinized based on Cohen’s kappa statistic. This statistical measure assesses the level of agreement between human annotators and automated sentiment classifiers,

Journal of Information System Research (JOSH)
Volume 5, No. 4, Juli 2024, pp 1445−1454
ISSN 2686-228X (media online)
https://ejurnal.seminar-id.com/index.php/josh/
DOI 10.47065/josh.v5i4.5501

providing a quantitative indicator of the reliability and consistency of sentiment classification results. This research ascertains the degree of agreement between the automated classifiers and human annotations by calculating the kappa statistic for both VADER and TextBlob across the dataset derived from viewer reviews of travel vlogs to Malaysia. This evaluation process enables a rigorous assessment of the effectiveness and validity of each sentiment analysis model, facilitating informed decision-making regarding suitability for sentiment analysis tasks in tourism-related contexts.

2.2.5 Deployment

During the deployment phase, the performance of the relevant model with the dataset was assessed, enabling the recommendation of content development for travel vlogs that elicit positive responses from viewers. By leveraging the insights gained from evaluating sentiment analysis models, content creators and tourism stakeholders identify effective strategies for crafting engaging and resonant travel vlog content. Moreover, deploying sentiment analysis models facilitates real-time monitoring and analysis of viewer feedback, allowing for agile adjustments and optimizations to enhance audience satisfaction and engagement. Consequently, this deployment phase is a pivotal stage in the iterative process of content development and audience engagement, enabling the creation of travel vlogs that effectively captivate and resonate with viewers, maximizing impact and reach.

3. RESULT AND DISCUSSION

The benefits of analyzing the distribution of polarity values derived from VADER and TextBlob in sentiment classification are multifaceted. Firstly, this analysis provides insights into the overall sentiment trends within the dataset, allowing for a comprehensive understanding of viewers' prevailing attitudes and emotions towards the subject matter. Additionally, by examining the distribution of polarity values, this research identifies patterns, outliers, and anomalies that may signify specific themes, topics, or sentiments prevalent in the data. Consequently, this analytical approach enables nuanced interpretation of sentiment dynamics, facilitating informed decision-making processes in content development, marketing strategies, and audience engagement initiatives.

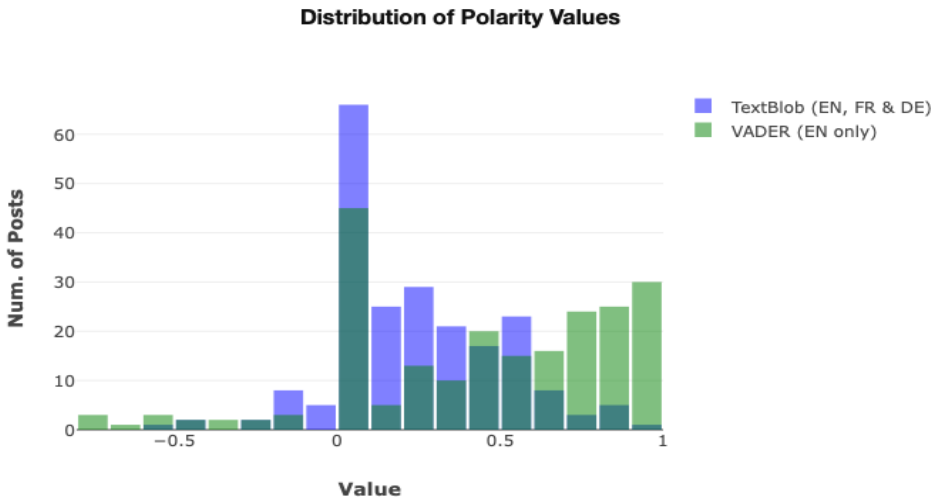


Figure 6. Distribution of Polarity Values based on VADER and TextBlob (First Video)

Figure 6 shows the distribution of polarity values based on VADER and TextBlob (first video). In the context of the first video, the distribution of sentiment polarity values reveals that out of 219 English language posts, 7 (4.40%) are categorized as having negative sentiments (polarity scores <= -0.05), 24 (15.09%) as neutral sentiments (polarity scores between -0.05 and 0.05), and 128 (80.50%) as positive sentiments (polarity scores >= 0.05). Excluding duplicates such as reposts or retweets, VADER and TextBlob exhibit agreement in categorizing 159 (72.60%) posts. This level of agreement, as determined by Cohen's kappa statistic of 0.393, is considered fair. This analysis underscores the utility of both VADER and TextBlob in sentiment classification tasks, albeit with varying degrees of agreement, thereby providing valuable insights for understanding audience sentiments and informing content development strategies.

Based on the distribution of polarity values derived from VADER and TextBlob in the first video, it is evident that VADER and TextBlob exhibit differences in sentiment classification outcomes. VADER predominantly classifies posts as positive sentiment, with 74.89% of posts falling within this category, followed by 17.81% classified as neutral sentiment and 7.31% as negative sentiment. Conversely, TextBlob categorizes a higher proportion of posts as neutral sentiment (26.48%) compared to VADER while also classifying 66.21% of posts as positive sentiment and 7.31% as negative sentiment. Notably, the TextBlob model trained on French language data classifies the sole post as neutral sentiment. This analysis underscores the importance of considering

Journal of Information System Research (JOSH)
Volume 5, No. 4, Juli 2024, pp 1445−1454
ISSN 2686-228X (media online)
https://ejurnal.seminar-id.com/index.php/josh/
DOI 10.47065/josh.v5i4.5501

the choice of sentiment analysis tool and language model in accurately capturing and interpreting sentiment dynamics within textual data, informing decision-making processes in content development and audience engagement strategies.

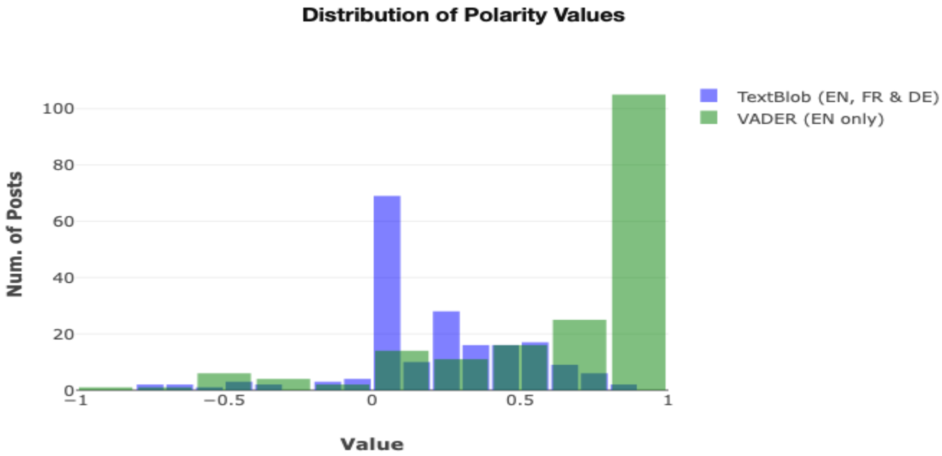


Figure 7. Distribution of Polarity Values based on VADER and TextBlob (Second Video)

Figure 7 shows the distribution of polarity values based on VADER and TextBlob (second video). In the context of the second video, the distribution of sentiment polarity values illustrates that out of 177 English language posts, 7 (6.09%) are classified as having negative sentiments (polarity scores <= -0.05), 6 (5.22%) as neutral sentiments (polarity scores between -0.05 and 0.05), and 102 (88.70%) as positive sentiments (polarity scores >= 0.05). Both VADER and TextBlob demonstrate agreement in categorizing 115 (64.97%) posts by excluding duplicates such as reposts or retweets. This level of agreement, as determined by Cohen's kappa statistic of 0.236, is considered fair. The findings underscore the applicability of VADER and TextBlob in discerning sentiment in textual data, albeit with some variance in agreement levels, thereby contributing valuable insights for understanding audience perceptions and guiding content creation strategies.

Based on the distribution of polarity values from VADER and TextBlob in the first video, discernible patterns emerge, highlighting variations in sentiment classification outcomes across different models and languages. VADER predominantly categorizes posts as positive sentiment, constituting 85.41% of the total posts, with a smaller proportion classified as negative (7.57%) or neutral (7.03%). In contrast, TextBlob demonstrates a more balanced distribution, with 58.92% of posts classified as positive, 33.51% as neutral, and 7.57% as negative. Furthermore, the language-specific TextBlob models for French and German exhibit unique patterns, with French posts predominantly classified as neutral sentiment (80.00%) and German posts entirely classified as neutral sentiment (100.00%). This analysis underscores the importance of considering both the sentiment analysis tool and language model in accurately interpreting sentiment dynamics within textual data, facilitating informed decision-making in content development and audience engagement strategies.

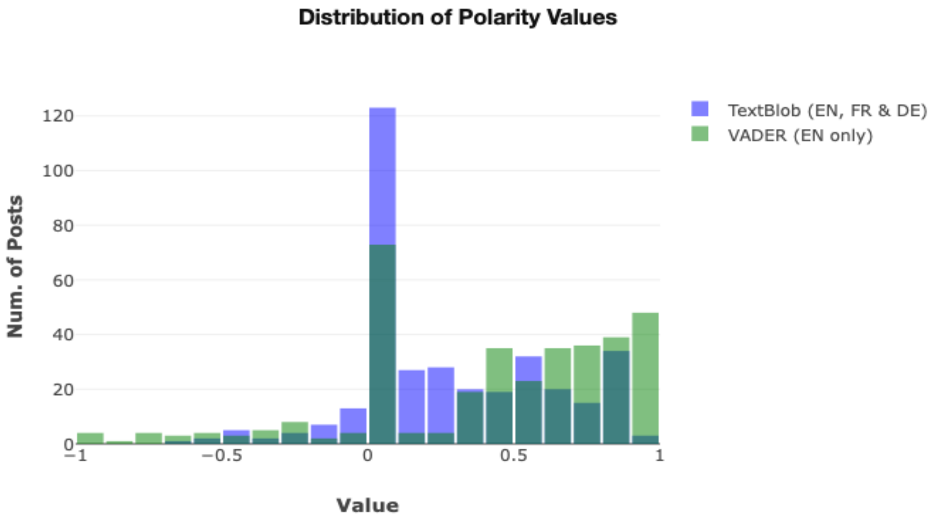


Figure 7. Distribution of Polarity Values based on VADER and TextBlob (Third Video)

Figure 7 shows the distribution of polarity values based on VADER and TextBlob (third video). In the context of the third video, the distribution of sentiment polarity values reveals that out of 351 English language posts, 14 (5.86%) are identified as having negative sentiments (polarity scores <= -0.05), 43 (17.99%) as neutral

Journal of Information System Research (JOSH)
Volume 5, No. 4, Juli 2024, pp 1445−1454
ISSN 2686-228X (media online)
https://ejurnal.seminar-id.com/index.php/josh/
DOI 10.47065/josh.v5i4.5501

sentiments (polarity scores between -0.05 and 0.05), and 182 (76.15%) as positive sentiments (polarity scores >= 0.05). Excluding duplicates such as reposts or retweets, VADER and TextBlob demonstrate agreement in categorizing 239 (68.09%) posts. This level of agreement, as indicated by Cohen's kappa statistic of 0.372, is considered fair. These findings underscore the reliability of both VADER and TextBlob in discerning sentiment in textual data, contributing valuable insights for understanding audience perceptions and guiding content creation strategies.

Based on the distribution of polarity values derived from VADER and TextBlob in the first video, distinct patterns emerge, revealing variations in sentiment classification outcomes across different models and languages. VADER's analysis indicates a predominant presence of positive sentiment, comprising 69.49% of the total posts, with smaller proportions classified as negative (10.45%) or neutral (20.06%). Conversely, TextBlob exhibits a more balanced distribution, with 60.45% of posts categorized as positive, 31.92% as neutral, and 7.63% as negative. Furthermore, the language-specific TextBlob model for French indicates a unique pattern, with all posts classified as neutral sentiment (100.00%). This analysis underscores the significance of considering both the sentiment analysis tool and language model in accurately interpreting sentiment dynamics within textual data, thus providing valuable insights for content development and audience engagement strategies.

4. CONCLUSION

In conclusion, the analysis of sentiment classification across three videos using VADER and TextBlob indicates notable patterns in sentiment distribution and agreement between the two models. Across the videos, VADER and TextBlob generally exhibit a fair level of agreement, ranging from 64.97% to 72.60%, as indicated by Cohen's kappa statistic. Moreover, the distribution of polarity values showcases varying proportions of negative, neutral, and positive sentiments across the videos, highlighting the dynamic nature of audience perceptions. These findings underscore the significance of employing multiple sentiment analysis tools and considering language-specific models for accurate sentiment classification in textual data. This research provides valuable insights for content creators, marketers, and analysts in understanding audience sentiments and tailoring content strategies accordingly.

ACKNOWLEDGMENT

I want to express my sincere gratitude to the Tourism Department, Faculty of Business Administration and Communication, Atma Jaya Catholic University of Indonesia, PUSDIPAR, and the LPPM (Lembaga Penelitian dan Pengabdian kepada Masyarakat) for the support in the publication of this research.

REFERENCES

[1] N. Dhariwal, S. C. Akunuri, Shivama, and K. S. Banu, "Audio and Text Sentiment Analysis of Radio Broadcasts," IEEE Access, vol. 11, no. November, pp. 126900–126916, 2023, doi: 10.1109/ACCESS.2023.3331226.

[2] T. Fontes, F. Murcos, E. Carneiro, J. Ribeiro, and R. J. F. Rossetti, "Leveraging Social Media as a Source of Mobility Intelligence: An NLP-Based Approach," IEEE Open J. Intell. Transp. Syst., vol. 4, no. September, pp. 663–681, 2023, doi: 10.1109/OJITS.2023.3308210.

[3] S. Bengesi, T. Oladunni, R. Olusegun, and H. Audu, "A Machine Learning-Sentiment Analysis on Monkeypox Outbreak: An Extensive Dataset to Show the Polarity of Public Opinion From Twitter Tweets," IEEE Access, vol. 11, no. February, pp. 11811–11826, 2023, doi: 10.1109/ACCESS.2023.3242290.

[4] E. Saad et al., "Determining the Efficiency of Drugs under Special Conditions from Users' Reviews on Healthcare Web Forums," IEEE Access, vol. 9, pp. 85721–85737, 2021, doi: 10.1109/ACCESS.2021.3088838.

[5] A. Mardjo and C. Choksuchat, "HyVADRF: Hybrid VADER-Random Forest and GWO for Bitcoin Tweet Sentiment Analysis," IEEE Access, vol. 10, no. September, pp. 101889–101897, 2022, doi: 10.1109/ACCESS.2022.3209662.

[6] F. Es-Sabery, I. Es-Sabery, A. Hair, B. Sainz-de-Abajo, and B. Garcia-Zapirain, "Emotion processing by applying a fuzzy-based Vader lexicon and a parallel deep belief network over massive data," IEEE Access, vol. 10, no. July, pp. 87870–87899, 2022, doi: 10.1109/ACCESS.2022.3200389.

[7] M. F. A. Gadi and M. A. Sicilia, "Annotators' Selection Impact on the Creation of a Sentiment Corpus for the Cryptocurrency Financial Domain," IEEE Access, vol. 11, no. November, pp. 131081–131088, 2023, doi: 10.1109/ACCESS.2023.3334260.

[8] O. Oyebode, R. Lomotey, and R. Orji, "'I Tried to Breastfeed but...': Exploring Factors Influencing Breastfeeding Behaviours Based on Tweets Using Machine Learning and Thematic Analysis," IEEE Access, vol. 9, pp. 61074–61089, 2021, doi: 10.1109/ACCESS.2021.3073079.

[9] R. K. Lomotey, S. Kumi, M. Hilton, R. Orji, and R. Deters, "Using Machine Learning to Establish the Concerns of Persons With HIV/AIDS During the COVID-19 Pandemic From Their Tweets," IEEE Access, vol. 11, no. April, pp. 37570–37601, 2023, doi: 10.1109/ACCESS.2023.3267050.

[10] D. Amangeldi, A. Usmanova, and P. Shamoi, "Understanding Environmental Posts: Sentiment and Emotion Analysis of Social Media Data," IEEE Access, vol. 12, no. March, pp. 33504–33523, 2024, doi: 10.1109/ACCESS.2024.3371585.

[11] A. Mahmoudi, D. Jemielniak, and L. Ciechanowski, "Assessing Accuracy: A Study of Lexicon and Rule-Based Packages in R and Python for Sentiment Analysis," IEEE Access, vol. 12, no. November 2023, pp. 20169–20180, 2024, doi: 10.1109/ACCESS.2024.3353692.

[12] M. Zubair, J. Ali, M. Alhussein, S. Hassan, K. Aurangzeb, and M. Umair, "An Improved Machine Learning-Driven Framework for Cryptocurrencies Price Prediction with Sentimental Cautioning," IEEE Access, vol. 12, no. February, pp. 51395–51418, 2024, doi: 10.1109/ACCESS.2024.3367129.
[13] I. Awajan, M. Mohamad, and A. Al-Quran, "Sentiment Analysis Technique and Neutrosophic Set Theory for Mining and Ranking Big Data from Online Reviews," IEEE Access, vol. 9, pp. 47338–47353, 2021, doi: 10.1109/ACCESS.2021.3067844.
[14] H. Kim and G. Qin, "Summarizing Students' Free Responses for an Introductory Algebra-Based Physics Course Survey Using Cluster and Sentiment Analysis," IEEE Access, vol. 11, no. July, pp. 89052–89066, 2023, doi: 10.1109/ACCESS.2023.3305260.
[15] F. T. Giuntini et al., "Modeling and Assessing the Temporal Behavior of Emotional and Depressive User Interactions on Social Networks," IEEE Access, vol. 9, pp. 93182–93194, 2021, doi: 10.1109/ACCESS.2021.3091801.
[16] S. K. Nayak, M. Garanayak, S. K. Swain, S. K. Panda, and D. Godavarthi, "An Intelligent Disease Prediction and Drug Recommendation Prototype by Using Multiple Approaches of Machine Learning Algorithms," IEEE Access, vol. 11, no. September, pp. 99304–99318, 2023, doi: 10.1109/ACCESS.2023.3314332.
[17] K. Aurangzeb, N. Ayub, and M. Alhussein, "Aspect Based Multi-Labeling Using SVM Based Ensembler," IEEE Access, vol. 9, pp. 26026–26040, 2021, doi: 10.1109/ACCESS.2021.3055768.
[18] A. Sittar, D. Major, C. Mello, D. Mladenic, and M. Grobelnik, "Political and Economic Patterns in COVID-19 News: From Lockdown to Vaccination," IEEE Access, vol. 10, pp. 40036–40050, 2022, doi: 10.1109/ACCESS.2022.3164692.
[19] L. Ketsbaia, B. Issac, X. Chen, and S. M. Jacob, "A Multi-Stage Machine Learning and Fuzzy Approach to Cyber-Hate Detection," IEEE Access, vol. 11, no. June, pp. 56046–56065, 2023, doi: 10.1109/ACCESS.2023.3282834.
[20] R. Parekh et al., "DL-GuesS: Deep Learning and Sentiment Analysis-Based Cryptocurrency Price Prediction," IEEE Access, vol. 10, no. April 2021, pp. 35398–35409, 2022, doi: 10.1109/ACCESS.2022.3163305.
[21] X. Zhou, J. W. C. Wong, D. Xie, R. Liang, and L. Huang, "What does the audience care? The effects of travel vlog information quality on travel intention," Total Qual. Manag. Bus. Excell., vol. 34, no. 15–16, pp. 2201–2219, 2023, doi: 10.1080/14783363.2023.2246908.
[22] S. K. Behera, P. Gautam, and S. K. Lenka, "Purchasing behavior of tourists toward prominent souvenirs of Odisha," Vilakshan - XIMB J. Manag., vol. 19, no. 2, pp. 165–176, Jan. 2022, doi: 10.1108/xjm-10-2020-0165.
[23] 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.
[24] T. T. Trinh Ngo, L. A. Phuong Bui, T. H. Dung Pham, N. Tram Mai, and K. Luan Bui, "Social media research trends in tourism city context," Int. J. Tour. Cities, vol. 8, no. 3, pp. 779–798, Jan. 2022, doi: 10.1108/IJTC-07-2021-0151.
[25] A. A. Arifiyanti, M. F. Pandji, and B. Utomo, "Analisis Sentimen Ulasan Pengunjung Objek Wisata Gunung Bromo pada Situs Tripadvisor," Explor. J. Sist. Inf. dan Telemat., vol. 13, no. 1, p. 32, 2022, doi: 10.36448/jsit.v13i1.2539.
[26] N. L. P. Merawati, A. Z. Amrullah, and Ismarmiaty, "Analisis Sentimen dan Pemodelan Topik Pariwisata Lombok Menggunakan Algoritma Naive Bayes dan Latent Dirichlet Allocation," Rekayasa Sist. dan Teknol. Inf., vol. 1, no. 10, pp. 123–131, 2021, doi: 10.29207/resti.v5i1.2587.
[27] Q. Su, A. Namin, and S. Ketron, "The effect of online company responses on app review quality," J. Consum. Mark., vol. 41, no. 1, pp. 110–125, Jan. 2024, doi: 10.1108/JCM-06-2023-6098.
[28] M. Lievonen, J. Bowden, and V. Luoma-aho, "Towards a typology of negative engagement behavior in social media," Serv. Ind. J., vol. 43, no. 3–4, pp. 238–259, 2023, doi: 10.1080/02642069.2022.2121961.
[29] P. Kumar, J. M. Mishra, and Y. V. Rao, "Analysing tourism destination promotion through Facebook by Destination Marketing Organizations of India," Curr. Issues Tour., vol. 25, no. 9, pp. 1416–1431, 2022, doi: 10.1080/13683500.2021.1921713.
[30] V. H. Luong, A. Manthiou, J. Kang, and C. Nguyen, "The building blocks of regenerative tourism and hospitality: a text-mining approach," Curr. Issues Tour., vol. 27, no. 3, pp. 361–380, 2024, doi: 10.1080/13683500.2023.2228974.