**Journal of Computer System and Informatics (JoSYC)**
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 5, No. 4, August 2024, Page 987-1000
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v5i4.5757

# Enhancing Tourism Digital Content Engagement through Sentiment and Toxicity Analysis: Application of Perspective, Vader, and TextBlob Models

**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
Submitted: 08/08/2024; Accepted: 25/08/2024; Published: 26/08/2024

**Abstract**−This research examines the engagement with tourism digital content for Sumba Island through sentiment and toxicity analysis. The study uses advanced models such as Perspective, Vader, and TextBlob to reveal an average toxicity score of 0.04066, indicating minimal harmful language. Sentiment classification shows a predominantly positive reception, with VADER identifying 81.69% positive, 12.96% neutral, and 5.35% negative sentiments. TextBlob analysis supports these findings, confirming the robustness of the sentiment evaluation. The research underscores the effectiveness of well-crafted digital content in promoting positive user engagement while maintaining low toxicity. The urgency of this research is emphasized by the increasing reliance on digital platforms for tourism marketing, where understanding audience perception is crucial for effective strategy development. The study employs the Digital Content Reviews and Analysis Framework, which ensures systematic data processing and comprehensive evaluation. This framework includes data cleansing, sentiment, toxicity scoring, and rigorous evaluation using multiple analytical models to enhance the reliability and applicability of the findings. Future recommendations include expanding the analysis to encompass visual content and non-English comments and incorporating advanced multimodal techniques to capture a holistic view of digital content engagement. Addressing these areas will further enrich the understanding and impact of tourism digital content, driving more effective and engaging marketing strategies in the competitive digital landscape.

**Keywords**: Tourism Digital; Content Engagement; Toxicity Score; Sentiment Classification; Sumba

## 1. INTRODUCTION

Engagement with tourism digital content presents a significant challenge in marketing destinations to attract domestic and international visitors. Effective digital content must capture the essence of the destination and resonate with diverse audiences through visually appealing and informative materials [1]–[4]. Enhancing this engagement involves leveraging advanced technologies such as augmented reality and interactive platforms, which enrich user experiences and facilitate virtual exploration of destinations [5]–[8]. This strategic approach increases visibility and fosters a deeper connection with potential tourists, ultimately driving higher visitor numbers. Consequently, a robust digital engagement strategy becomes imperative for destination marketing, ensuring sustained interest and competitiveness in the global tourism market.

Toxicity and sentiment analysis represent relevant approaches for evaluating the engagement of digital content intended to encourage tourist visits to published destinations. These analytical methods provide insights into public perceptions and emotional reactions, crucial for refining content to ensure it resonates positively with the target audience [9], [10]. By identifying and mitigating negative sentiments and enhancing positive feedback, these analyses help optimize the appeal and effectiveness of promotional materials [11]–[13]. This strategic application not only aids in creating more compelling digital content but also enhances the likelihood of successful tourist engagement, ultimately driving higher visitation rates to the featured destinations.

The Perspective model is employed to identify toxicity scores, while Vader and TextBlob are utilized for sentiment classification. Perspective analyzes textual data to detect and measure toxic language, providing a quantitative toxicity score that aids in monitoring and mitigating harmful content [14], [15]. Concurrently, Vader and TextBlob offer robust sentiment classification capabilities, with Vader excelling in analyzing social media text and TextBlob being advantageous for its simplicity and adaptability [16]. Employing these models in tandem enables a comprehensive analysis of digital content, ensuring both the identification of toxic language and the accurate classification of sentiment, thereby enhancing online interactions' overall quality and relevance.

This research aims to identify viewer responses to digital content through comments and to calculate textual data using the Perspective model to determine Toxicity, Severe Toxicity, Identity Attack, Insult, Profanity, and Threat. By leveraging the Perspective model, the study quantifies various forms of toxic language, providing a detailed analysis of negative interactions. Vader and TextBlob models also classify sentiments into positive, neutral, and negative categories, offering a nuanced understanding of audience sentiment. This dual-method approach enhances the accuracy of content evaluation and provides valuable insights for improving digital content engagement and moderation strategies.

The urgency of this research lies in its potential to address the escalating issues of digital toxicity and sentiment misclassification in online platforms. As digital interactions increasingly dominate social and commercial spheres, accurately identifying and mitigating harmful content becomes paramount [17]–[19]. Employing advanced analytical models such as Perspective, Vader, and TextBlob provides a systematic

Journal of Computer System and Informatics (JoSYC)
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 5, No. 4, August 2024, Page 987-1000
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v5i4.5757

approach to evaluating and enhancing the quality of digital engagement [20]–[22]. This research is essential for improving user experience and safeguarding online communities against the detrimental effects of toxic interactions. Consequently, the study's findings will be pivotal in guiding future content moderation and digital marketing strategies, ensuring a more positive and secure digital environment.

This research's theoretical and practical implications are multifaceted, providing significant advancements in academic and applied contexts. Theoretically, this study enriches the existing body of knowledge by integrating sophisticated models like Perspective, Vader, and TextBlob to analyze digital content, thus offering new insights into online toxicity and sentiment dynamics. Practically, applying these models in real-world scenarios enhances the capability of digital platforms to manage user interactions more effectively, promoting healthier online environments. This research contributes to academic discourse and provides actionable strategies for improving digital content management and user engagement by bridging the gap between theoretical exploration and practical application.

Similar research in digital content analysis has employed various methodologies to examine toxicity and sentiment in online interactions. Studies utilizing models such as BERT and RoBERTa have demonstrated the efficacy of advanced natural language processing techniques in detecting harmful content and sentiment patterns [23]–[27]. These investigations highlight the importance of integrating multiple analytical frameworks to comprehensively understand digital engagement [28]–[31]. By comparing the effectiveness of different models, these studies contribute to optimizing content moderation strategies and enhancing user experience. Thus, the convergence of findings from similar research underscores the necessity for continuous innovation in the analysis of digital content.

The limitation of this research lies in the framework employed, precisely the Digital Content Reviews and Analysis Framework, which predominantly emphasizes contextual analysis aligned with content. While adept at providing in-depth contextual insights, this framework may overlook specific quantitative metrics crucial for a comprehensive evaluation of digital engagement [32]–[36]. Additionally, its focus on content-specific context might limit the generalizability of findings across diverse digital platforms [37]–[39]. Therefore, while the framework offers valuable context-based analysis, incorporating a more holistic approach could enhance the robustness and applicability of the research outcomes. Consequently, future studies might benefit from integrating complementary analytical models to address these limitations.

The recommendation for further research emphasizes the need to integrate more diverse analytical frameworks to enhance the comprehensiveness of digital content evaluation. Future studies should explore incorporating quantitative metrics alongside contextual analysis to provide a more balanced and robust assessment of digital engagement. Additionally, examining the applicability of various natural language processing models across different digital platforms could yield valuable insights into the generalizability of findings. By adopting a multifaceted approach, subsequent research can address the limitations of current frameworks, thereby contributing to the development of more effective strategies for managing digital content and improving user interactions.

## 2. RESEARCH METHODOLOGY

### 2.1 Gap Analysis

The research gap identified in previous studies highlights the lack of comprehensive frameworks that simultaneously address toxicity and sentiment analysis in tourism digital content. Prior research often focused on either sentiment analysis or toxicity detection in isolation, failing to provide a holistic view of digital engagement. This study, titled "Tourism Digital Content Engagement through Toxicity and Sentiment Analysis: Perspective, Vader, and TextBlob," introduces a novel approach by integrating these analyses to offer a more nuanced understanding of user interactions with tourism content. This dual-faceted methodology is an existing research gap, and the precision and effectiveness of digital content evaluation contribute significantly to tourism marketing and online content moderation.

The novelty of this research lies in its integrative approach to analyzing tourism digital content engagement through both toxicity and sentiment analysis, employing advanced models such as Perspective, Vader, and TextBlob [40], [40], [40]–[42], [42]. This study uniquely combines the detection of sentiment classification of sentiments, providing a comprehensive framework that addresses the multifaceted nature of digital interactions. By leveraging these models, the research offers an innovative methodology that enhances the accuracy and depth of content evaluation. This dual-faceted analysis fills a critical gap in the existing literature and sets a new standard for future digital content engagement and tourism marketing studies.

The importance of this research lies in its ability to provide critical insights into the effectiveness of tourism digital content through advanced sentiment and toxicity analysis. By employing sophisticated models such as Perspective, Vader, and TextBlob, the study offers a comprehensive understanding of audience engagement and sentiment, essential for developing targeted marketing strategies. This research highlights the positive impact of well-crafted digital content and identifies areas needing improvement, enhancing the overall

Journal of Computer System and Informatics (JoSYC)
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 5, No. 4, August 2024, Page 987-1000
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v5i4.5757

user experience. Ultimately, the findings of this study are pivotal for informing future tourism marketing efforts, ensuring that digital content effectively captures and retains the interest of potential tourists.

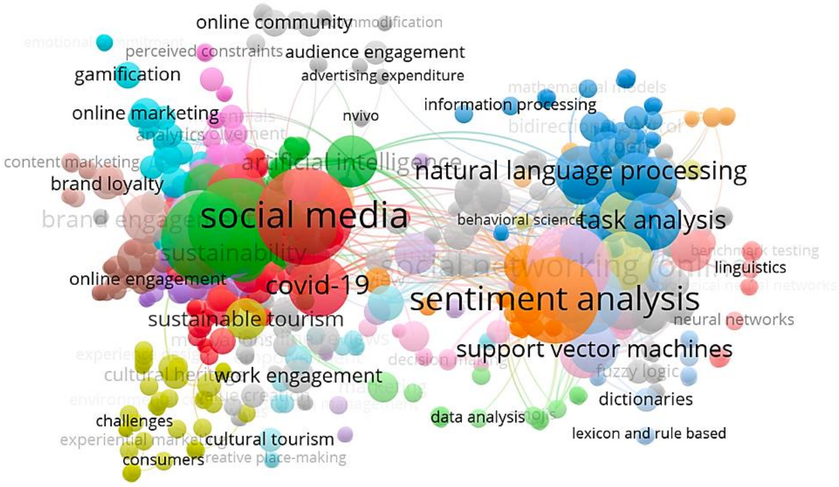



**Figure 1.** Network Visualization

Figure 1 shows network visualization using Vosviewer. The current trends in research indicate a significant focus on social media, sentiment analysis, and sustainable tourism, reflecting the evolving landscape of digital engagement and marketing. Integrating artificial intelligence and natural language processing has become paramount, enabling more sophisticated and nuanced analysis of user interactions. Moreover, the impact of COVID-19 has intensified the need for understanding online community behavior and virtual engagement, driving researchers to explore innovative solutions for maintaining brand loyalty and audience engagement in a digital-first world. These trends underscore the critical importance of interdisciplinary approaches, combining technology, behavioral science, and marketing strategies to adapt to and leverage the rapidly changing digital environment for more effective and sustainable tourism outcomes.

The importance of using this framework lies in its ability to systematically and comprehensively analyze digital content, ensuring high accuracy and reliability in the findings. By integrating multiple advanced analytical models such as Perspective, Vader, and TextBlob, the framework allows for a nuanced understanding of sentiment and toxicity in user interactions [43], [44]. This multifaceted approach enhances the robustness of the analysis and provides actionable insights for optimizing digital content. This framework ensures that tourism marketing strategies are informed by precise and detailed data, leading to more effective engagement with potential tourists and a better overall user experience.

### 2.2 Digital Content Reviews and Analysis Framework

This research employs the Digital Content Reviews and Analysis Framework, which provides a robust structure for examining and interpreting digital content. By focusing on contextual analysis, this framework enables a nuanced understanding of the intricate dynamics present within user interactions and content engagement. The methodology encompasses various dimensions of content evaluation, including toxicity detection and sentiment classification, thereby ensuring a comprehensive assessment [41], [44], [44]. Utilizing this framework significantly enhances the analytical precision and depth of the study, offering valuable insights into digital engagement patterns. Consequently, this approach advances academic understanding and informs practical strategies for effective digital content management.

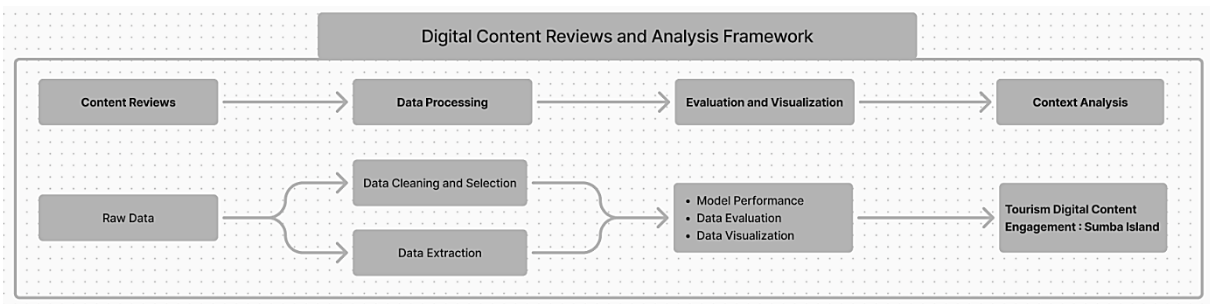



**Figure 2.** Digital Content Reviews and Analysis Framework

Figure 2 shows the digital content reviews and analysis framework. The Digital Content Reviews and Analysis Framework comprises several systematic stages to ensure thorough and accurate content evaluation. Initially, raw data undergoes data cleansing and selection, followed by data extraction to prepare it for detailed

Journal of Computer System and Informatics (JoSYC)
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 5, No. 4, August 2024, Page 987-1000
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v5i4.5757

analysis. The data processing stage then facilitates the transformation and structuring of the data, which is essential for subsequent evaluation and visualization. This methodology includes assessing model performance and visualizing data insights, culminating in a comprehensive context analysis. Ultimately, this structured approach enhances the depth and accuracy of digital content reviews and provides actionable insights, particularly in the tourism digital content engagement case study on Sumba Island.

The comment data to be processed is sourced from the YouTube video "SUMBA ISLAND, Indonesia - The Secret Paradise," which has garnered 558 comments and 561,432 views since its upload on August 10, 2019. This substantial volume of user-generated content offers a rich dataset for analysis, providing insights into viewer engagement and sentiment towards the depicted destination. Analyzing these comments through advanced models such as Perspective, Vader, and TextBlob will allow for a detailed understanding of positive and negative sentiments and the prevalence of toxic language. Consequently, this analysis will contribute significantly to developing more effective digital marketing strategies for tourism on Sumba Island.

### 2.2.1 Content Reviews

The initial stage involves identifying the content based on the post-per-day statistics of the content. This metric provides a quantitative measure of content frequency, crucial for understanding post-temporal distribution and consistency. By analyzing the post-per-day statistics, patterns in content delivery can be discerned, revealing insights into the engagement strategies employed. This approach not only aids in assessing the effectiveness of content scheduling but also informs the optimization of future content dissemination plans. Leveraging this statistical analysis ensures a more strategic and impactful digital content management process.

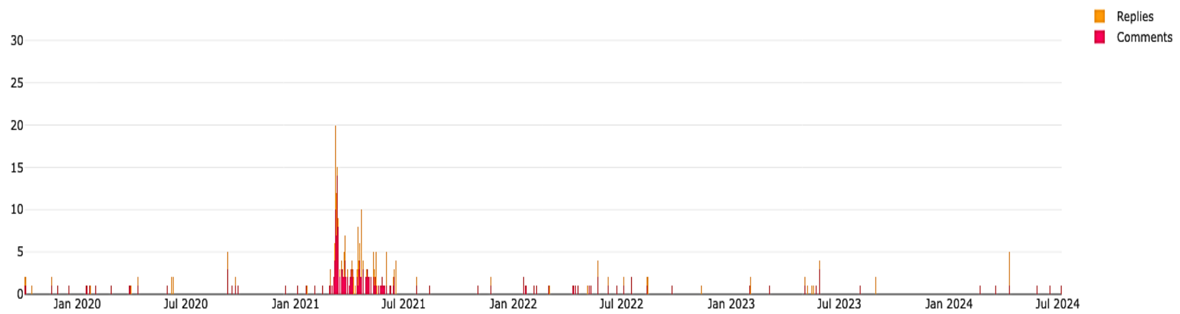



**Figure 3.** Post-Per-Day Statistic of the Content

Figure 3 shows the post-per-day statistic. The initial phase of this data analysis involves identifying content based on the post-per-day statistic, which provides a detailed overview of user interaction trends over time. The provided data highlights fluctuations in comment activity, with notable spikes in early 2021, indicating periods of heightened engagement. These patterns are essential for understanding the temporal dynamics of content interaction, suggesting that certain events or promotional activities may have influenced increased viewer participation. By leveraging this temporal analysis, more strategic content planning and audience engagement techniques can be developed, optimizing digital marketing efforts and enhancing user interaction.

The data illustrates the temporal dynamics of digital content engagement for the "SUMBA ISLAND, Indonesia - The Secret Paradise" video on YouTube. A notable surge in comments and replies occurred in early 2021, with the highest activity observed in March and April 2021. This peak suggests a significant event or effective promotional strategy during this period that captured the audience's attention and spurred interaction. Such fluctuations in engagement metrics underscore the importance of timing in digital content strategies. The spikes indicate heightened viewer interest, which could be correlated with external factors such as travel trends, media coverage, or seasonal promotions. Understanding these patterns allows for more informed decisions in content scheduling and marketing efforts.

Furthermore, the gradual decline in comments post-mid-2021 implies that maintaining consistent engagement requires continuous innovation in content delivery. Regular monitoring and analysis of post-per-day statistics enable content creators to identify the most influential periods for audience engagement and adjust their strategies accordingly. In conclusion, this temporal analysis of digital content engagement highlights the critical role of strategic timing and adaptive content strategies in sustaining viewer interaction and interest over time.

Based on the data from the top-ten posters, it is evident that @romanbader is the most active contributor, with 126 posts constituting 66.3% of the total contributions. @TIMOR_RAYA follows with 33 posts, accounting for 17.4%, while the remaining contributors, including @jessicabella116, @no-sparringholloway, and others, have significantly fewer posts. This distribution highlights a predominant engagement from a single user, which suggests a potential influence or interest focus on the content by @romanbader. This disproportionate activity could skew the overall sentiment and engagement metrics, indicating the need for a more balanced interaction from a broader audience to achieve a comprehensive

Journal of Computer System and Informatics (JoSYC)
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 5, No. 4, August 2024, Page 987-1000
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v5i4.5757

analysis. Consequently, strategies to diversify engagement and encourage broader participation should be considered to enhance the reliability and representativeness of the content evaluation.


**Figure 4.** Top Ten Poster

Figure 4 shows the top ten posters of the content. The data visualization of the top-ten posters on the "SUMBA ISLAND, Indonesia - The Secret Paradise" video reveals significant insights into user engagement patterns. The leading contributor, @romanbader, with 126 posts, dominates the interaction, comprising 66.3% of the total comments and replies. This overwhelming participation indicates this user's solid interest or potential promotional intent, which can significantly influence the overall engagement dynamics. The second highest contributor, @TIMOR_RAYA, with 33 posts (17.4%), also shows a notable activity level, though it is considerably lower than that of @romanbader. The remaining eight contributors, including @jessicabella116 (7 posts), @no-sparringholloway (6 posts), and others with even fewer posts, collectively comprise a minor proportion of the engagement.

This disproportionate contribution from a small number of users suggests that the engagement metrics for this video are not evenly distributed across a broad audience. Such a concentration can lead to a biased representation of sentiment and engagement levels, where the views and interactions of a few individuals overshadow the broader audience's reactions. From a strategic perspective, this analysis underscores the importance of fostering more balanced and widespread user interaction. Encouraging a diverse range of viewers to engage with the content would provide a more accurate and representative understanding of the audience's sentiment and engagement. This could be achieved through targeted outreach, engaging content strategies, and interactive campaigns designed to attract a broader spectrum of viewers. In conclusion, while the high engagement from top contributors like @romanbader highlights individual solid interest, a more diversified interaction is essential for reliable and comprehensive content evaluation. Balancing engagement would enhance the quality of insights derived from the data, ultimately leading to more effective digital marketing strategies.

### 2.2.2 Data Processing: Cleaning, Selection and Extraction

Following the comprehension of the data context, data cleansing is meticulously executed using the RapidMiner application. The procedure commences with tokenization, segmenting the text into individual tokens, and is followed by case transformation to ensure uniformity across the dataset. Subsequently, token filtering based on length and removing stopwords are implemented to discard common yet non-informative words. The final phase involves applying the Porter stemming algorithm to reduce words to their root forms. This systematic methodology significantly enhances data quality, ensuring that ensuing analyses are predicated on purified and pertinent data. Ultimately, utilizing RapidMiner for data cleansing amplifies the precision and efficacy of the data analysis process.

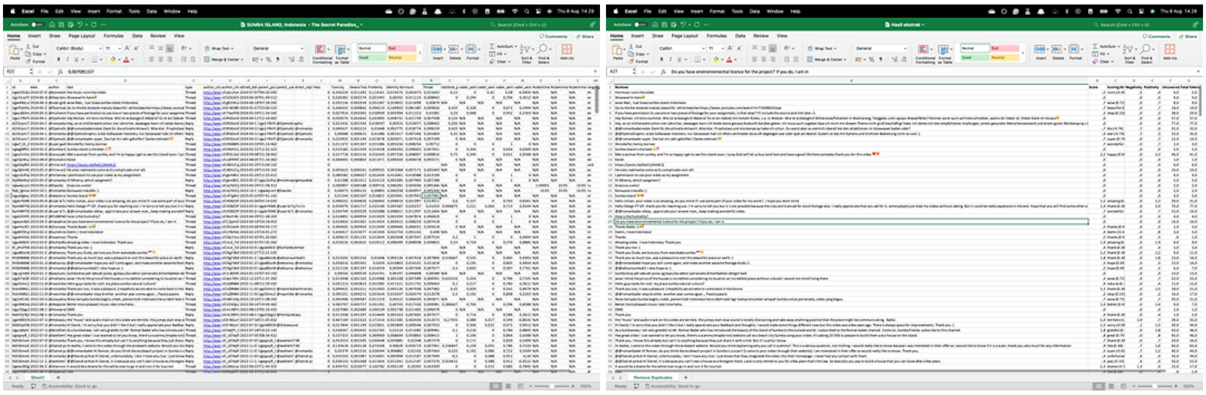

**Figure 5.** Data Cleaning Process

Figure 5 shows the cleaning process. The cleaned text data comprises 559 entries processed through sophisticated operations. Initially, the tokenize operator segments the text into discrete tokens, ensuring each word is individually analyzed. It is followed by the transform cases operator, which standardizes the text to a uniform case, enhancing consistency. Subsequent steps involve filtering tokens by length and removing stopwords, eliminating common but uninformative words. The final operation employs the stemming technique to reduce words to their root forms. This comprehensive data cleansing process ensures the resulting dataset is refined and relevant, optimizing subsequent analyses' accuracy and effectiveness.

The results of the data cleansing process reveal that the cleaned text data amounts to 509 entries, necessitating further selection for data processing purposes. This refined dataset represents a significant reduction from the original volume, indicating the effective removal of extraneous and non-informative

Journal of Computer System and Informatics (JoSYC)
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 5, No. 4, August 2024, Page 987-1000
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v5i4.5757

elements. The subsequent step involves a careful selection process to ensure the remaining data is relevant and valuable for the intended analysis. This meticulous approach is crucial for enhancing the quality and reliability of the data, thereby ensuring that the final analysis is both accurate and meaningful. Consequently, this selection phase is integral to the overall success of the data processing workflow.



(a) (b)

**Figure 6.** Data Selection Process (a and b)

Figure 6 shows the data selection process. The selected data comprises 509 text entries with polynomial and binomial types, ensuring a structured approach to data classification. Sentiment classification into negative and positive categories is meticulously aligned with the extracted scores using the Vader model. This model provides a nuanced analysis of sentiment intensity, enabling precise categorization of text data based on sentiment scores. The systematic application of Vader ensures that each entry is accurately classified, thereby enhancing the reliability of the sentiment analysis. Ultimately, this approach guarantees that the dataset is thoroughly evaluated, providing robust insights for subsequent data processing and interpretation.

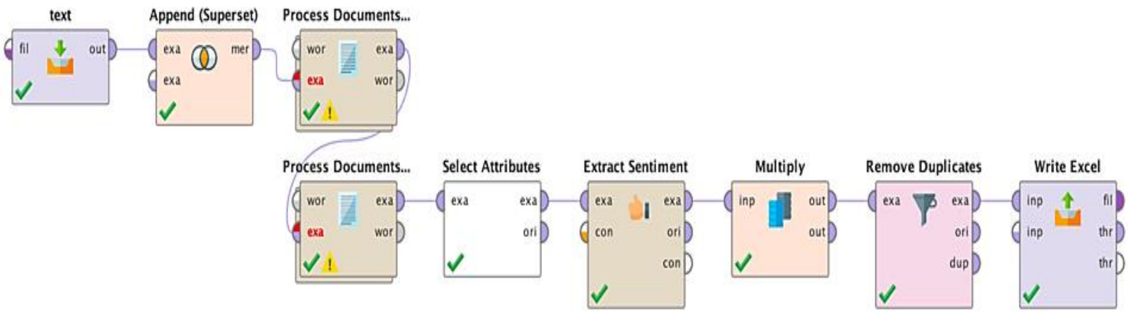



**Figure 7.** Extraction Process Using Vader Operator in Rapidminer

Figure 7 shows the extraction process using the Vader operator in Rapidminer. The sentiment extraction process utilizing Vader follows a systematic and comprehensive workflow. Initially, the text data undergoes preparation and is appended with necessary attributes for further analysis. The document processing stage involves transforming the raw text into a structured format ready for attribute selection. Subsequently, the Vader model extracts sentiment scores, categorizing the data into positive, neutral, and negative sentiments. This step is crucial for understanding the emotional tone of the text entries. The processed data is then subjected to multiplication and duplication removal to ensure accuracy and consistency. Finally, the refined data is exported into an Excel file for detailed analysis and reporting. This meticulous process ensures the sentiment analysis is thorough and reliable, providing valuable insights for subsequent evaluations.

### 2.2.3 Data Evaluation and Visualization

The sentiment extraction and classification results are subsequently evaluated using the NBC, DT, k-NN, and SVM algorithms with SMOTE. These machine-learning algorithms are initially applied to assess the accuracy and robustness of the sentiment classifications derived from the Vader model. Implementing SMOTE addresses class imbalances, ensuring a more equitable evaluation across different sentiment categories. Employing a diverse set of algorithms provides a comprehensive analysis, highlighting the strengths and limitations of each method in sentiment classification. Ultimately, this rigorous evaluation process enhances the reliability and validity of the sentiment analysis, ensuring robust and actionable insights.

**Journal of Computer System and Informatics (JoSYC)**
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 5, No. 4, August 2024, Page 987-1000
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v5i4.5757

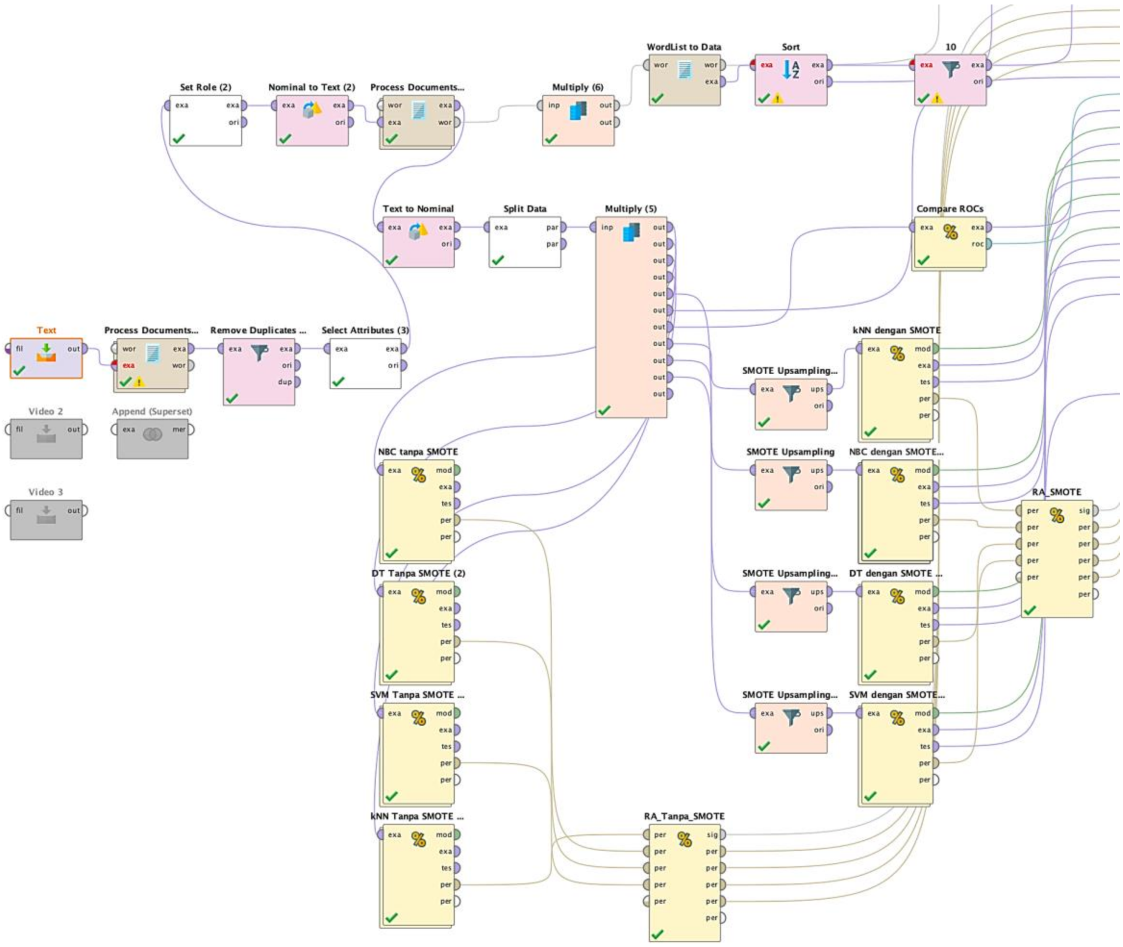



**Figure 8.** Model Evaluation

Figure 8 shows the model evaluation using NBC, k-NN, DT, and SVM. The algorithm performance evaluation reveals an accuracy of 76.55% with a standard deviation of 5.06%, as indicated by the micro average. The confusion matrix shows that the model correctly identified 188 negative and 331 positive instances, with minimal misclassifications. The area under the curve (AUC) values are impressive, with an optimistic AUC of 0.972, a general AUC of 0.970, and a pessimistic AUC of 0.969, reflecting the model's high discriminatory power. Precision stands at 68.95%, suggesting room for improvement in minimizing false positives, while the recall rate of 97.63% underscores the model's efficacy in capturing true positive sentiments. The f-measure, combining precision and recall, is robust at 80.73%, demonstrating a well-balanced performance across all metrics. This comprehensive evaluation underscores the model's substantial potential for reliable sentiment classification while highlighting areas for refinement to enhance overall accuracy.

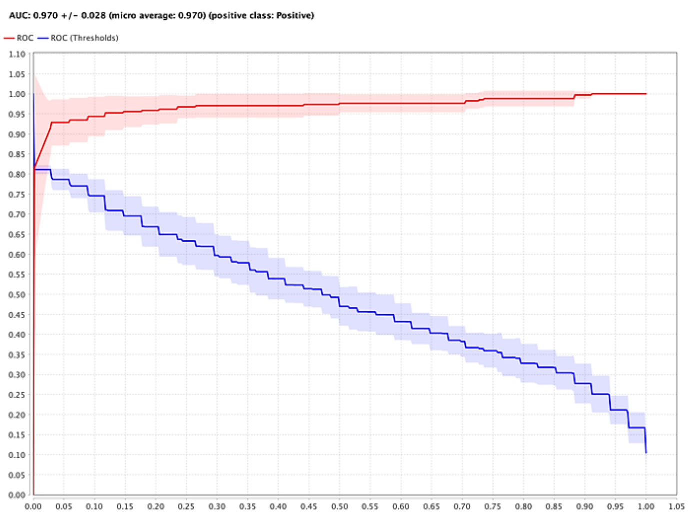



**Figure 9.** Area Under Curve (AUC) of SVM with SMOTE.

ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 5, No. 4, August 2024, Page 987-1000
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v5i4.5757

Figure 9 shows SVM's Area Under the Curve (AUC) with SMOTE. The visualization of the Area Under Curve (AUC) for the SVM model with SMOTE illustrates the model's robust performance in distinguishing between positive and negative sentiment classes. The AUC value of 0.970, with a micro average of 0.970 and a standard deviation of 0.028, underscores the high discriminatory power of the model. The Receiver Operating Characteristic (ROC) curve in red for the positive class demonstrates a steep ascent toward the top-left corner, indicating a high valid positive rate and a low false positive rate. This graphical representation confirms the model's effectiveness in accurately classifying sentiments, benefiting from the balanced dataset provided by SMOTE. Ultimately, the high AUC value reflects the SVM model's reliability and precision in sentiment analysis, making it a valuable tool for extracting meaningful insights from text data.

Visualizing the Area Under Curve (AUC) for the SVM model with SMOTE provides crucial insights into tourism digital content engagement. With an impressive AUC value of 0.970, the model accurately distinguishes between positive and negative sentiments in user comments related to tourism content. This high level of precision is essential for understanding how digital content influences tourist perceptions and engagement. The ROC curve's sharp ascent towards the top-left corner indicates that the model has a high valid positive rate and a low false positive rate. In tourism, the model can reliably identify genuinely positive feedback from tourists, which is invaluable for destination marketing and content strategy. Accurate sentiment classification allows marketers to pinpoint which aspects of their content resonate most with viewers, enabling them to refine their messaging and enhance overall engagement.

Furthermore, using SMOTE to balance the dataset ensures that the model's performance is not biased towards more prevalent classes, such as positive or neutral sentiments, which are standard in promotional tourism content. This balance is critical for maintaining the integrity of the analysis, ensuring that both positive and negative sentiments are accurately represented and addressed. In conclusion, the high AUC value and the robust performance of the SVM model with SMOTE indicate that it is a reliable tool for analyzing sentiment in tourism digital content. This capability allows for more effective content strategies tailored to elicit positive engagement and enhance the tourist experience, ultimately driving higher satisfaction and loyalty.

### 2.2.4 Context Analysis: Tourism Digital Content Engagement

The importance of context analysis for tourism digital content engagement through video cannot be overstated. Context analysis provides a deeper understanding of how and why certain content resonates with audiences by examining the surrounding factors that influence viewer reactions. This approach involves evaluating various elements such as cultural relevance, seasonal trends, and the emotional tone of the video, which collectively impact viewer engagement. Analyzing these contextual factors enables marketers to tailor their content more effectively, ensuring it appeals to the target audience's interests and preferences. Consequently, context analysis is essential for optimizing video content to enhance viewer interaction, drive engagement, and, ultimately, promote tourism destinations more successfully.

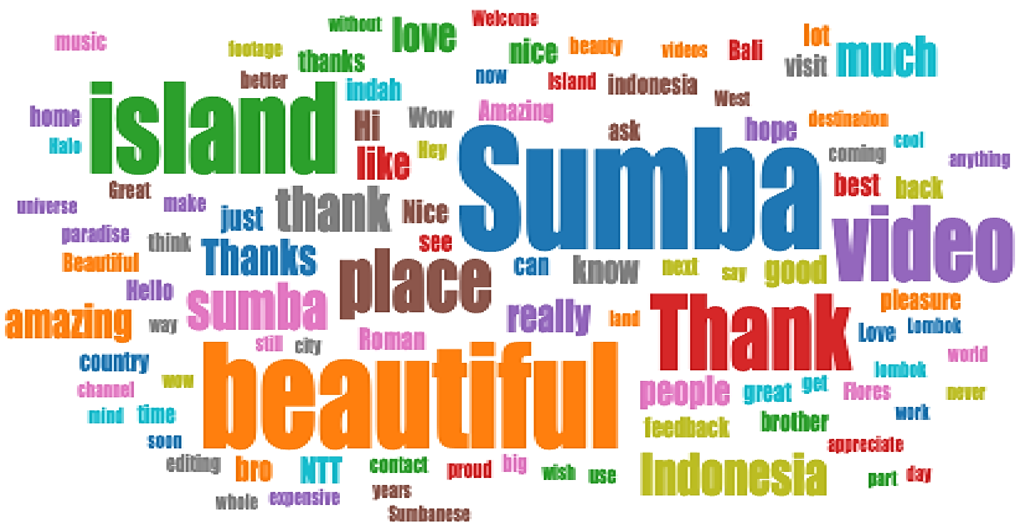



**Figure 10.** Wordclouds of Dataset

Figure 10 shows the word clouds of the dataset. Based on the word cloud, the most frequently mentioned topics in viewer comments on the video are prominently highlighted. Key terms such as "Sumba," "beautiful," "island," and "thank" indicate a strong appreciation and admiration for the beauty of Sumba Island, suggesting that viewers are highly engaged with the aesthetic and visual appeal of the content. Additionally, words like "Indonesia," "place," and "amazing" further emphasize the positive sentiment towards the destination. This analysis underscores the video's success in capturing and conveying the allure of Sumba, effectively engaging the audience and evoking a favorable response. Consequently, the frequent occurrence of these terms illustrates

Journal of Computer System and Informatics (JoSYC)
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 5, No. 4, August 2024, Page 987-1000
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v5i4.5757

the video's impact on viewers, reinforcing the importance of visually compelling content in promoting tourism destinations.

The analysis and interpretation of viewer comments within the context of Tourism Digital Content Engagement reveal significant insights into audience interaction and sentiment. Frequent mentions of words such as "beautiful," "island," and "Sumba" indicate a strong positive reception, highlighting the visual and aesthetic appeal of the destination. The recurring appreciation expressed through terms like "thank," "amazing," and "nice" suggests that the content successfully evokes emotional responses and gratitude from viewers. This engagement is critical for tourism marketing, as positive viewer sentiment can enhance destination appeal and drive potential tourist interest. Therefore, the effective use of captivating visuals and emotive language in digital content is essential for maximizing engagement and fostering a favorable perception of tourism destinations.

# 3. RESULT AND DISCUSSION

The discussion in this research is divided into two sections: Tourism Digital Content Engagement through Toxicity and Sentiment Analysis: Perspective, Vader, and TextBlob, and the general discussion. The first section focuses on applying advanced analytical models to evaluate the engagement levels of digital content by assessing toxicity and sentiment. These models provide a comprehensive understanding of user interactions, offering valuable insights into how digital content is perceived. The subsequent discussion section synthesizes these findings, exploring their implications for tourism marketing strategies and content creation. Ultimately, this bifurcated approach ensures a thorough examination of digital content engagement, facilitating a deeper understanding of its impact on tourism.

## 3.1 Tourism Digital Content Engagement through Toxicity and Sentiment Analysis: Perspective, Vader, and TextBlob

Tourism Digital Content Engagement through Toxicity and Sentiment Analysis employs advanced analytical models such as Perspective, Vader, and TextBlob to evaluate user interactions with digital content comprehensively. Perspective focuses on identifying and quantifying toxic language, providing insights into the negative aspects of user comments. Vader and TextBlob, on the other hand, are utilized to classify sentiments into positive, neutral, and negative categories, offering a nuanced understanding of the emotional tone conveyed by viewers. The integration of these models allows for a detailed analysis of both the positive engagement and potential negative impacts, thus enabling more effective and strategic digital content creation. Ultimately, this approach enhances the understanding of audience reception and optimizes tourism marketing efforts.

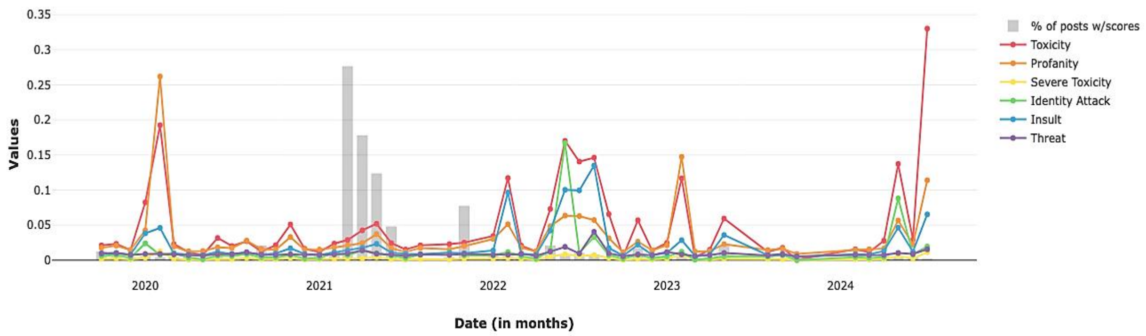



**Figure 11.** Average Toxicity Score Per-Month

Figure 11 shows the average toxicity score per month. Communalytic's analysis of 477 posts out of 558 using the Perspective API reveals various negative sentiment indicators. The average toxicity score is 0.04066, with a peak of 0.47120, indicating occasional high toxicity. Severe toxicity remains relatively low, with an average of 0.00276 and a maximum of 0.09386. Identity attacks and insults present slightly higher averages of 0.01019 and 0.01928, respectively, with their maxima approaching 0.49314 and 0.49627, reflecting frequent personal attacks. Profanity, with an average score of 0.02719 and a maximum of 0.54193, underscores the regular use of offensive language. While less common, threats have an average score of 0.01027 and a peak of 0.28362. These findings highlight various forms of negative engagement, necessitating ongoing monitoring and mitigation strategies to foster a positive online environment.

The average monthly toxicity score, as depicted in the graph, reveals significant fluctuations over the observed period. The scores show occasional spikes, with notable peaks in early 2020, mid-2021, and late 2023, indicating periods of heightened negative sentiment or harmful language in user comments. These variations suggest that certain events or content releases may have triggered more toxic interactions, highlighting the importance of contextual analysis. Understanding these patterns is crucial for developing strategies to mitigate

Journal of Computer System and Informatics (JoSYC)
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 5, No. 4, August 2024, Page 987-1000
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v5i4.5757

negative engagement and enhance user experience. Ultimately, monitoring and addressing monthly toxicity scores can significantly improve the quality and impact of tourism digital content.

Based on the 378 out of 558 posts analysis, sentiment distribution varies significantly across different languages and models. The VADER model in English analyzed 355 posts, revealing that 81.69% exhibited positive sentiment, 12.96% were neutral, and 5.35% were negative. In contrast, the TextBlob model in English showed a slightly lower positive sentiment at 72.11%, with 23.66% neutral and 4.23% negative. All sentiments were neutral for the French posts analyzed with TextBlob, accounting for 100%. The German TextBlob analysis demonstrated that 41.18% of posts were positive and 58.82% were neutral, with no negative sentiments recorded. These results suggest that while positive sentiment predominates, neutrality and negativity can vary depending on the language and model used, highlighting the importance of utilizing multiple analytical tools for a comprehensive sentiment analysis.

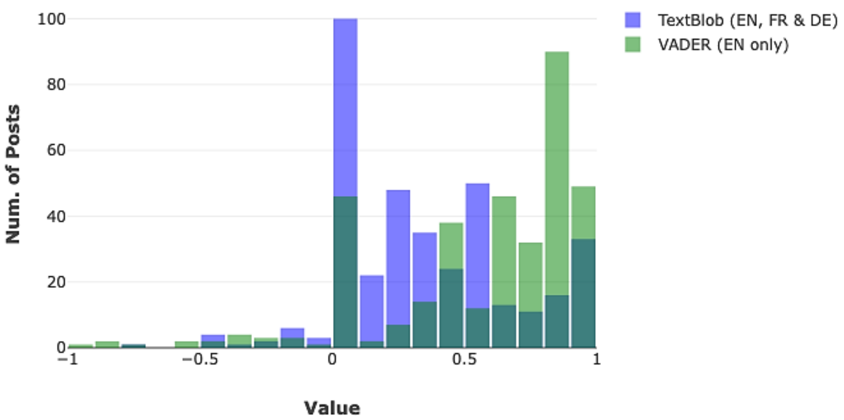



**Figure 12.** Distribution of Polarity Value

Figure 12 shows the distribution of polarity value. Excluding duplicates such as reposts and retweets, VADER and TextBlob agree on categorizing 249 out of 312 English language posts, representing a 79.81% agreement level. This substantial concurrence between the two sentiment analysis models indicates a moderate level of reliability, as reflected by a Cohen's kappa statistic of 0.506. Such an agreement is noteworthy, considering the inherent differences in the algorithms' approaches to sentiment classification. This moderate level of agreement underscores the value of using multiple models to achieve a more comprehensive sentiment analysis, enhancing the robustness and accuracy of the overall findings.

Expressly, both libraries agree on the categorization of sentiments in the analyzed posts as follows: 9 posts (3.61%) exhibit negative sentiments with polarity scores less than or equal to -0.05, 35 posts (14.06%) reflect neutral sentiments with polarity scores between -0.05 and 0.05, and 205 posts (82.33%) display positive sentiments with polarity scores greater than or equal to 0.05. This agreement highlights the predominant positive sentiment expressed by the majority of users. The concordance in sentiment classification between the two libraries underscores their reliability in identifying sentiment trends accurately. Consequently, this alignment enhances confidence in the analytical results, providing a robust foundation for further sentiment analysis in digital content engagement.

The analysis and interpretation of sentiment agreement between VADER and TextBlob reveal significant insights into the nature of user engagement with digital content. Specifically, the two libraries concur that 3.61% of posts exhibit negative sentiments, 14.06% are neutral, and a substantial 82.33% display positive sentiments. This overwhelming positive sentiment underscores a generally favorable reception of the content, suggesting that most users have a positive experience or opinion. The relatively low percentage of negative sentiments indicates that the content is well-received, with minimal adverse reactions. The moderate agreement on neutral sentiments reflects a balanced view among a smaller subset of users, who neither strongly favor nor oppose the content.

The high level of agreement between VADER and TextBlob, particularly in identifying positive sentiments, enhances the reliability of the sentiment analysis. This concordance suggests that both models effectively capture the emotional tone conveyed in user posts. The robust identification of sentiment trends is crucial for understanding user engagement and guiding content creators and marketers in tailoring their strategies to reinforce positive aspects and address negative feedback. In conclusion, the alignment in sentiment analysis between VADER and TextBlob confirms the predominance of positive engagement, providing valuable insights for optimizing digital content to enhance user experience and satisfaction. This analysis is a critical tool for improving content strategy and fostering a more engaging and positive interaction with the audience.

### 3.2 Discussion: Tourism Content Engagement

Tourism content engagement is a critical factor in the success of digital marketing strategies promoting travel destinations. Analyzing user interactions with digital content, mainly through sentiment and toxicity

**Journal of Computer System and Informatics (JoSYC)**
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 5, No. 4, August 2024, Page 987-1000
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v5i4.5757

assessments, provides valuable insights into audience perceptions and preferences [45], [46]. Positive engagement, as indicated by the high percentage of favorable sentiments, underscores the effectiveness of well-crafted visual and narrative elements in capturing audience interest [47]–[49]. Conversely, understanding the instances and sources of negative sentiment allows for targeted improvements and mitigation strategies, ensuring a more balanced and appealing content presentation. Ultimately, leveraging detailed engagement analysis fosters a deeper connection with potential tourists, enhancing the reach and impact of tourism marketing efforts.

Tourism digital content of Sumba Island plays a pivotal role in showcasing its unique attractions and captivating landscapes to a global audience. High-quality visual and narrative elements in the content significantly enhance viewer engagement, effectively highlighting the island's natural beauty and cultural richness. Analyzing viewer interactions and sentiment towards this content reveals a predominantly positive reception, indicating successful communication of the island's appeal [50]. This positive engagement boosts the island's visibility and encourages potential tourists to explore Sumba further. Consequently, well-executed digital content is a powerful tool in promoting Sumba Island as a premier travel destination, driving the region's tourism growth and economic benefits.

Aspects to be highlighted in creating destination marketing content are essential for attracting and engaging potential tourists. Firstly, the content should emphasize the destination's unique attractions and cultural heritage, offering a compelling narrative that sets it apart from other locations. Additionally, high-quality visuals and multimedia elements should be incorporated to visually captivate the audience and provide a virtual experience of the destination [49], [51]–[54]. Engaging storytelling that includes testimonials and experiences from past visitors can also enhance credibility and relatability. A strategic focus on these aspects maximizes the destination's appeal and fosters a deeper emotional connection with the audience, driving higher engagement and tourism interest.

The role of digital technology in creating destination marketing content is pivotal for enhancing tourism promotion and engagement. Digital tools enable the production of high-quality visuals and interactive media that vividly showcase a destination's attractions and cultural assets. Advanced analytics and data-driven insights facilitate the customization of marketing strategies, ensuring that content resonates with target audiences [55], [56]. Additionally, virtual and augmented reality technologies provide immersive experiences, allowing potential tourists to explore destinations virtually. Ultimately, leveraging digital technology amplifies the reach and effectiveness of marketing campaigns and fosters a deeper connection with prospective travelers, significantly boosting tourism interest and engagement.

The relevance of this research lies in its comprehensive analysis of tourism digital content engagement through advanced sentiment and toxicity analysis. By employing models such as Perspective, Vader, and TextBlob, the study provides critical insights into how digital content is perceived by audiences, thereby informing more effective marketing strategies. This approach aligns with current trends in leveraging big data and machine learning to enhance content effectiveness and user engagement. Furthermore, the research's focus on Sumba Island offers valuable case-specific findings that can be generalized to other tourism destinations. Ultimately, this research contributes significantly to the field of tourism marketing, providing actionable insights for optimizing digital content to attract and retain tourists.

The limitations of this research primarily stem from the reliance on specific sentiment and toxicity analysis models, which may not capture the full complexity of user interactions and sentiments. The study's focus on textual data excludes other forms of digital content, such as images and videos, which are crucial in tourism marketing. Additionally, the scope is limited to the analysis of comments in English, potentially overlooking significant insights from non-English speaking audiences. Future research should consider incorporating multimodal analysis, including visual content, and expanding the linguistic range to include a more diverse set of user inputs. Addressing these limitations will enhance the comprehensiveness and applicability of the findings, providing deeper insights into digital content engagement across varied platforms and cultural contexts.

# 4. CONCLUSION

The research concludes that engagement with tourism digital content for Sumba Island demonstrates a predominantly positive sentiment with relatively low toxicity levels. The analysis, using Perspective, Vader, and TextBlob models, revealed that the average toxicity score of user comments was 0.04066, indicating minimal harmful language. Sentiment classification results showed that most posts were positive, with VADER identifying 81.69% positive, 12.96% neutral, and 5.35% negative sentiments. TextBlob analysis corroborated these findings, although with slightly different proportions, highlighting the robustness of the sentiment analysis. These results underscore the effectiveness of well-crafted digital content in promoting positive user engagement while maintaining a low-toxicity environment. The increasing reliance on digital platforms for tourism marketing underscores the urgency of this research. Understanding how audiences perceive and engage with digital content is crucial for developing effective marketing strategies that attract and retain tourists in a highly competitive digital landscape. Employing the Digital Content Reviews and Analysis Framework, this study systematically processed and evaluated the data, ensuring a thorough and accurate analysis. This framework included data cleansing, sentiment, toxicity scoring, and rigorous evaluation using multiple analytical models. This structured approach not only enhances the reliability of the findings but also provides a replicable methodology for future research in digital content engagement. Future recommendations include expanding the scope of analysis to encompass visual content and non-English comments and incorporating more advanced multimodal analysis techniques to capture a comprehensive view of digital content engagement. Addressing these areas will further enhance the understanding and impact of tourism digital content, ultimately driving more effective and engaging marketing strategies.

## 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*).

## REFERENCES

[1] B. Garner and D. Kim, "Analyzing user-generated content to improve customer satisfaction at local wine tourism destinations: an analysis of Yelp and TripAdvisor reviews," *Consum. Behav. Tour. Hosp.*, vol. 17, no. 4, pp. 413–435, Jan. 2022, doi: 10.1108/CBTH-03-2022-0077.

[2] R. Correia, M. Cunha, A. Carvalho, and B. Sousa, "The (wasted) potential of digital communication in low-density destinations: the case of three Portuguese municipalities," *J. Hosp. Tour. Insights*, vol. 7, no. 2, pp. 723–742, Jan. 2024, doi: 10.1108/JHTI-09-2022-0422.

[3] T. Chauhan, S. Sindhu, and R. S. Mor, "Modelling the factors impacting customer engagement for branded content in healthcare," *Int. J. Pharm. Healthc. Mark.*, vol. 18, no. 1, pp. 102–121, Jan. 2024, doi: 10.1108/IJPHM-03-2022-0030.

[4] J. Bowden and A. Mirzaei, "Consumer engagement within retail communication channels: an examination of online brand communities and digital content marketing initiatives," *Eur. J. Mark.*, vol. 55, no. 5, pp. 1411–1439, Jan. 2020, doi: 10.1108/EJM-01-2018-0007.

[5] L. Florido-Benítez, "The impact of tourism promotion in tourist destinations: a bibliometric study," *Int. J. Tour. Cities*, vol. 8, no. 4, pp. 844–882, Jan. 2022, doi: 10.1108/IJTC-09-2021-0191.

[6] M. Mariani, S. Bresciani, and G. B. Dagnino, "The competitive productivity (CP) of tourism destinations: an integrative conceptual framework and a reflection on big data and analytics," *Int. J. Contemp. Hosp. Manag.*, vol. 33, no. 9, pp. 2970–3002, Jan. 2021, doi: 10.1108/IJCHM-09-2020-1102.

[7] E. K. K. Tsetse, M. A. Mahmoud, C. Blankson, and R. Odoom, "The impact of stakeholder market orientation on sustainability performance at tourism destinations," *Manag. Res. Rev.*, vol. 45, no. 7, pp. 929–955, Jan. 2022, doi: 10.1108/MRR-04-2021-0290.

[8] G. Tripathi and P. Wasan, "Positioning of tourist destinations in the digital era: a review of online customer feedback," *Worldw. Hosp. Tour. Themes*, vol. 13, no. 2, pp. 275–290, Jan. 2021, doi: 10.1108/WHATT-09-2020-0122.

[9] C. H. Perera, L. T. Van Nguyen, and R. Nayak, "Brand engagement on social media and its impact on brand equity in higher education: integrating the social identity perspective," *Int. J. Educ. Manag.*, vol. 37, no. 6–7, pp. 1335–1359, Jan. 2023, doi: 10.1108/IJEM-05-2023-0260.

[10] S. Bilderback, "Tides of change: employee training strategies for promoting risk-taking tourism in the North Sea," *Mar. Econ. Manag.*, vol. 7, no. 1, pp. 67–78, Jan. 2024, doi: 10.1108/maem-01-2024-0002.

[11] Y. D. Lestari, F. Saidah, and A. N. Aliya Putri, "Effect of destination competitiveness attributes on tourists' intention to visit halal tourism destination in Indonesia," *J. Islam. Mark.*, vol. 14, no. 4, pp. 937–965, Jan. 2023, doi: 10.1108/JIMA-12-2020-0368.

[12] L. J. Salangsang, M. J. Liwanag, and P. A. Notorio, "A content analysis of Asian countries' tourism video advertisements: a luxury travel perspective," *Consum. Behav. Tour. Hosp.*, vol. 17, no. 1, pp. 76–88, Jan. 2022, doi: 10.1108/CBTH-05-2021-0141.

[13] C. Ji, S. Mieiro, and G. Huang, "How social media advertising features influence consumption and sharing intentions: the mediation of customer engagement," *J. Res. Interact. Mark.*, vol. 16, no. 1, pp. 137–153, Jan. 2022, doi: 10.1108/JRIM-04-2020-0067.

[14] M. S. Viñán-Ludeña and L. M. de Campos, "Analyzing tourist data on Twitter: a case study in the province of Granada at Spain," *J. Hosp. Tour. Insights*, vol. 5, no. 2, pp. 435–464, Jan. 2022, doi: 10.1108/JHTI-11-2020-0209.

[15] S. M. C. Loureiro, R. G. Bilro, and A. Japutra, "The effect of consumer-generated media stimuli on emotions and consumer brand engagement," *J. Prod. Brand Manag.*, vol. 29, no. 3, pp. 387–408, Jan. 2020, doi: 10.1108/JPBM-11-2018-2120.

[16] 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.

[17] K. Taiminen and C. Ranaweera, "Fostering brand engagement and value-laden trusted B2B relationships through digital content marketing: The role of brand's helpfulness," *Eur. J. Mark.*, vol. 53, no. 9, pp. 1759–1781, Jan. 2019, doi: 10.1108/EJM-10-2017-0794.

[18] G. Taheri, F. Mohammadi, and M. Jami Pour, "What drives social co-creation in tourism? An empirical study," *Foresight*, vol. 26, no. 1, pp. 114–135, Jan. 2024, doi: 10.1108/FS-01-2023-0015.

[19] D. Buhalis, A. Papathanassis, and M. Vafeidou, "Smart cruising: smart technology applications and their diffusion in cruise tourism," *J. Hosp. Tour. Technol.*, vol. 13, no. 4, pp. 626–649, Jan. 2022, doi: 10.1108/JHTT-05-2021-0155.

[20] E. S. Valenzuela-Gálvez, Á. Garrido-Morgado, and Ó. González-Benito, "Boost your email marketing campaign!

Emojis as visual stimuli to influence customer engagement," *J. Res. Interact. Mark.*, vol. 17, no. 3, pp. 337–352, Jan. 2023, doi: 10.1108/JRIM-02-2021-0033.
[21] S. S. Matusse, X. Xi, and I. M. Joaquim, "Assessment of strategies to enhance the online presence of the Mozambican government website on tourism destination marketing," *Int. Trade, Polit. Dev.*, vol. 7, no. 1, pp. 16–35, Jan. 2023, doi: 10.1108/itpd-07-2022-0012.
[22] Y. Huang, J. Finsterwalder, N. (Chris) Chen, and F. R. L. Crawford, "Online student engagement and place attachment to campus in the new service marketplace: an exploratory study," *J. Serv. Mark.*, vol. 36, no. 4, pp. 597–611, Jan. 2022, doi: 10.1108/JSM-04-2021-0148.
[23] P. Rajapaksha, R. Farahbakhsh, and N. Crespi, "BERT, XLNet or RoBERTa: The Best Transfer Learning Model to Detect Clickbaits," *IEEE Access*, vol. 9, pp. 154704–154716, 2021, doi: 10.1109/ACCESS.2021.3128742.
[24] J.-H. Kim, S.-W. Park, J.-Y. Kim, J. Park, S.-H. Jung, and C.-B. Sim, "RoBERTa-CoA: RoBERTa-Based Effective Finetuning Method Using Co-Attention," *IEEE Access*, vol. 11, pp. 120292–120303, 2023, doi: 10.1109/ACCESS.2023.3328352.
[25] R. Anggrainingsih, G. M. Hassan, and A. Datta, "CE-BERT: Concise and Efficient BERT-Based Model for Detecting Rumors on Twitter," *IEEE Access*, vol. 11, pp. 80207–80217, 2023, doi: 10.1109/ACCESS.2023.3299858.
[26] K. L. Tan, C. P. Lee, K. S. M. Anbananthen, and K. M. Lim, "RoBERTa-LSTM: A Hybrid Model for Sentiment Analysis With Transformer and Recurrent Neural Network," *IEEE Access*, vol. 10, pp. 21517–21525, 2022, doi: 10.1109/ACCESS.2022.3152828.
[27] S. Salim and O. Lahcen, "A BERT-Enhanced Exploration of Web and Mobile Request Safety Through Advanced NLP Models and Hybrid Architectures," *IEEE Access*, vol. 12, pp. 76180–76193, 2024, doi: 10.1109/ACCESS.2024.3406413.
[28] J. Seo, S. Lee, L. Liu, and W. Choi, "TA-SBERT: Token Attention Sentence-BERT for Improving Sentence Representation," *IEEE Access*, vol. 10, pp. 39119–39128, 2022, doi: 10.1109/ACCESS.2022.3164769.
[29] P. Thiengburanathum and P. Charoenkwan, "SETAR: Stacking Ensemble Learning for Thai Sentiment Analysis Using RoBERTa and Hybrid Feature Representation," *IEEE Access*, vol. 11, pp. 92822–92837, 2023, doi: 10.1109/ACCESS.2023.3308951.
[30] A. S. Imran, H. Hodnefjeld, Z. Kastrati, N. Fatima, S. M. Daudpota, and M. A. Wani, "Classifying European Court of Human Rights Cases Using Transformer-Based Techniques," *IEEE Access*, vol. 11, pp. 55664–55676, 2023, doi: 10.1109/ACCESS.2023.3279034.
[31] H. Zhang, S. Sun, Y. Hu, J. Liu, and Y. Guo, "Sentiment Classification for Chinese Text Based on Interactive Multitask Learning," *IEEE Access*, vol. 8, pp. 129626–129635, 2020, doi: 10.1109/ACCESS.2020.3007889.
[32] A. Singh and S. Munjal, "How is the hospitality and tourism industry in India responding to the dynamic digital era?," *Worldw. Hosp. Tour. Themes*, vol. 13, no. 2, pp. 163–167, Jan. 2021, doi: 10.1108/WHATT-09-2020-0118.
[33] W. G. Ryan, A. Fenton, W. Ahmed, and P. Scarf, "Recognizing events 4.0: the digital maturity of events," *Int. J. Event Festiv. Manag.*, vol. 11, no. 1, pp. 47–68, Jan. 2020, doi: 10.1108/IJEFM-12-2019-0060.
[34] W. Tafesse and B. P. Wood, "Social media influencers' community and content strategy and follower engagement behavior in the presence of competition: an Instagram-based investigation," *J. Prod. Brand Manag.*, vol. 32, no. 3, pp. 406–419, Jan. 2023, doi: 10.1108/JPBM-02-2022-3851.
[35] P. Wang and Q. Huang, "Digital influencers, social power and consumer engagement in social commerce," *Internet Res.*, vol. 33, no. 1, pp. 178–207, Jan. 2023, doi: 10.1108/INTR-08-2020-0467.
[36] 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.
[37] A. Yousaf, I. Amin, D. Jaziri, and A. Mishra, "Effect of message orientation/vividness on consumer engagement for travel brands on social networking sites," *J. Prod. Brand Manag.*, vol. 30, no. 1, pp. 44–57, Jan. 2021, doi: 10.1108/JPBM-08-2019-2546.
[38] K. Yawised, D. Apasrawirote, M. Chatrangsan, and P. Muneesawang, "Travelling in the digital world: exploring the adoption of augmented reality (AR) through mobile application in hospitality business sector," *J. Adv. Manag. Res.*, vol. 20, no. 4, pp. 599–622, Jan. 2023, doi: 10.1108/JAMR-01-2023-0023.
[39] L. Ryan Bengtsson and J. Edlom, "Commodifying participation through choreographed engagement: the Taylor Swift case," *Arts Mark.*, vol. 13, no. 2, pp. 65–79, Jan. 2023, doi: 10.1108/AAM-07-2022-0034.
[40] 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.
[41] M. F. A. Gadi and M. Á. Sicilia, "Annotators' Selection Impact on the Creation of a Sentiment Corpus for the Cryptocurrency Financial Domain," *IEEE Access*, vol. 11, pp. 131081–131088, 2023, doi: 10.1109/ACCESS.2023.3334260.
[42] 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, pp. 37570–37601, 2023, doi: 10.1109/ACCESS.2023.3267050.
[43] M. Khalid *et al.*, "Novel Sentiment Majority Voting Classifier and Transfer Learning-Based Feature Engineering for Sentiment Analysis of Deepfake Tweets," *IEEE Access*, vol. 12, pp. 67117–67129, 2024, doi: 10.1109/ACCESS.2024.3398582.
[44] 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, pp. 11811–11826, 2023, doi: 10.1109/ACCESS.2023.3242290.
[45] S. Liasidou, G. Berjozkina, and K. Pipyros, "Mitigating seasonality in tourism by attracting young travellers: evidence from Cyprus," *Worldw. Hosp. Tour. Themes*, vol. 14, no. 5, pp. 461–469, Jan. 2022, doi: 10.1108/WHATT-07-2022-0077.
[46] A. F. Abbas, M. G. Khwaja, A. Z. Abbasi, and A. Hameed, "Market mavenism, tourists' co-creation experience, loyalty, vaxication intention: mediating role of travel incentives in the post-COVID-19 environment," *Consum. Behav.*

ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 5, No. 4, August 2024, Page 987-1000
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v5i4.5757

*Tour. Hosp.*, vol. 18, no. 4, pp. 532–550, Jan. 2023, doi: 10.1108/CBTH-01-2023-0007.

[47] V. Kaushal and R. Yadav, "Exploring luxury hospitality customer experience of Maldives tourists amidst COVID-19 pandemic," *Consum. Behav. Tour. Hosp.*, vol. 19, no. 1, pp. 140–165, Jan. 2024, doi: 10.1108/CBTH-04-2022-0085.

[48] E. Ballester, C. Ruiz-Mafé, and N. Rubio, "Females' customer engagement with eco-friendly restaurants in Instagram: the role of past visits," *Int. J. Contemp. Hosp. Manag.*, vol. 35, no. 6, pp. 2267–2288, Jan. 2023, doi: 10.1108/IJCHM-02-2022-0178.

[49] A. M. Morrison, S. Bag, and K. Mandal, "Virtual reality's impact on destination visit intentions and the moderating role of amateur photography," *Tour. Rev.*, vol. 79, no. 2, pp. 355–377, Jan. 2024, doi: 10.1108/TR-12-2022-0621.

[50] 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.

[51] T. Hassan and M. I. Saleh, "Tourism metaverse from the attribution theory lens: a metaverse behavioral map and future directions," *Tour. Rev.*, vol. 79, no. 5, pp. 1088–1104, Jan. 2024, doi: 10.1108/TR-07-2023-0516.

[52] J. Song, H. Moon, and M. Kim, "When do customers engage in brand pages? Effects of social presence," *Int. J. Contemp. Hosp. Manag.*, vol. 31, no. 9, pp. 3627–3645, Jan. 2019, doi: 10.1108/IJCHM-10-2018-0816.

[53] W. K. S. Leung, M. K. Chang, M. L. Cheung, and S. Shi, "VR tourism experiences and tourist behavior intention in COVID-19: an experience economy and mood management perspective," *Inf. Technol. People*, vol. 36, no. 3, pp. 1095–1125, Jan. 2023, doi: 10.1108/ITP-06-2021-0423.

[54] W. K. S. Leung, M. L. Cheung, M. K. Chang, S. Shi, S. Y. Tse, and L. Yusrini, "The role of virtual reality interactivity in building tourists' memorable experiences and post-adoption intentions in the COVID-19 era," *J. Hosp. Tour. Technol.*, vol. 13, no. 3, pp. 481–499, Jan. 2022, doi: 10.1108/JHTT-03-2021-0088.

[55] M. Simoni, A. Sorrentino, D. Leone, and A. Caporuscio, "Boosting the pre-purchase experience through virtual reality. Insights from the cruise industry," *J. Hosp. Tour. Technol.*, vol. 13, no. 1, pp. 140–156, Jan. 2022, doi: 10.1108/JHTT-09-2020-0243.

[56] D. Buhalis, M. S. Lin, and D. Leung, "Metaverse as a driver for customer experience and value co-creation: implications for hospitality and tourism management and marketing," *Int. J. Contemp. Hosp. Manag.*, vol. 35, no. 2, pp. 701–716, Jan. 2023, doi: 10.1108/IJCHM-05-2022-0631.