Journal of Information System Research (JOSH)
Volume 6, No. 1, Oktober 2024, pp 85−99
ISSN 2686-228X (media online)
https://ejurnal.seminar-id.com/index.php/josh/
DOI 10.47065/josh.v6i1.5961

Toxicity Score and Sentiment Classification of Backpacker Content Reviews using SVM enhanced by SMOTE

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: **24/09/2024;** Accepted: **09/10/2024;** Published: **14/10/2024**

Abstract−This research explores the dynamics of backpacker tourism in Indonesia by analyzing online content from various regions, including Bandung, Dieng, Borobudur, Ijen, Bromo, Tumpak Sewu, Malang, Banyuwangi, and Bali. Using the Digital Content Reviews and Analysis Framework, the study systematically processed user-generated content to assess sentiment and toxicity levels. The analysis revealed that while most interactions were non-toxic, there were occasional spikes in harmful language, particularly in the categories of profanity and identity attacks. For example, toxicity scores in Malang, Banyuwangi, and Bali averaged 0.06995, with peaks reaching 0.78207, underscoring the need for ongoing content moderation. In addition, the study employed a Support Vector Machine (SVM) model enhanced by SMOTE to handle class imbalance. The model achieved an accuracy of 82.64% and a recall rate of 97.39%, demonstrating its effectiveness in identifying positive cases with minimal false negatives. The AUC scores, ranging from 0.970 to 0.979, indicated strong discriminatory power. These findings highlight the potential of using machine learning models to analyze large-scale, imbalanced datasets in tourism-related research. Overall, this study provides valuable insights into traveler perceptions of Indonesia's backpacker destinations, emphasizing the importance of context in understanding online discourse. The integration of toxicity analysis and SVM modeling offers practical implications for improving tourism management, content moderation, and promoting sustainable tourism practices.

Keywords: Content Reviews; Sentiment Analysis; Toxicity Detection; Bacpacker Tourism

1. INTRODUCTION

Backpacking has emerged as a compelling form of tourism for adventurers, captivating individuals who seek more than the conventional travel experience. This activity not only emphasizes the exploration of remote or lesser-known destinations but also fosters an authentic connection with diverse cultures and environments [1]. As adventurers immerse themselves in these unique experiences, the content generated from such journeys becomes inherently engaging, offering a narrative rich in personal discovery and unfiltered exploration [2]. The appeal of this content lies in its spontaneity and raw documentation, contrasting with the highly curated nature of mainstream travel media [3]. Furthermore, backpackers often encounter unpredictable challenges and serendipitous moments, which add an element of intrigue and relatability to their stories. This dynamic nature of backpacking content creates a lasting impact on audiences, making it a source of inspiration for future travelers and a valuable contribution to the evolving discourse on experiential tourism. Ultimately, backpacking, as both an activity and a content generator, remains an influential component in shaping contemporary travel narratives.

Within the realm of tourism, backpacking has emerged as an appealing and cost-effective travel activity, offering substantial advantages for budget-conscious travelers. This form of tourism emphasizes independent exploration with minimal expenses, allowing participants to allocate resources efficiently while still experiencing diverse cultural and geographical landscapes [4]. The economic benefits of backpacking stem from the use of affordable accommodations, public transportation, and locally sourced services, which collectively reduce the overall financial burden of travel [5].

This financial accessibility not only democratizes travel by making it attainable for a broader demographic but also fosters a deeper, more immersive connection with local communities [6]. From an analytical perspective, the affordability and flexibility inherent in backpacking make it a preferred choice for many travelers, aligning with contemporary trends that prioritize authentic and sustainable travel experiences. In this regard, backpacking serves not only as an enriching travel activity but also as a practical approach to tourism planning.

This study aims to analyze the comparative performance of sentiment classification and toxicity score calculations for comments on backpacker content using the Perspective and VADER models. The investigation centers on evaluating the accuracy and reliability of these two distinct natural language processing tools in assessing both the emotional tone and the level of harmful content within user-generated comments [7]–[9]. Perspective, known for its advanced machine learning algorithms, provides a nuanced approach to toxicity detection, while VADER, a lexicon-based model, offers rapid and interpretable sentiment analysis [10]–[13]. It is argued that each model presents unique advantages, with Perspective excelling in identifying subtle toxic language and VADER being more efficient in capturing general sentiment trends [14]–[18]. Through an analytical comparison of results, this research seeks to highlight the strengths and limitations of both models, offering insights into their applicability in real-world content moderation. Ultimately, such findings contribute to the broader discourse on the optimization of automated sentiment analysis tools in the digital age.

Journal of Information System Research (JOSH)
Volume 6, No. 1, Oktober 2024, pp 85−99
ISSN 2686-228X (media online)
https://ejurnal.seminar-id.com/index.php/josh/
DOI 10.47065/josh.v6i1.5961

The urgency of this research lies in its potential to address critical gaps in understanding and managing the increasing complexity of online discourse, particularly in the context of user-generated content. As digital platforms continue to grow, the need for effective sentiment classification and toxicity detection has become paramount for ensuring healthy communication environments [19], [20]. This study contributes to this need by evaluating the effectiveness of advanced language models in accurately detecting and moderating harmful language. Given the rising incidents of online harassment, misinformation, and negative sentiment, such research holds significant importance for both academic and practical applications. The comparative analysis of sentiment models like Perspective and VADER offers a valuable framework for enhancing content moderation strategies, which is particularly relevant for platforms aiming to balance free expression with the prevention of harmful interactions [21]–[23]. Consequently, this research not only addresses immediate concerns in digital communication but also provides long-term insights into optimizing automated systems for safer and more inclusive online spaces.

The framework employed in this study is the Digital Content Reviews and Analysis Framework, which serves as a comprehensive tool for systematically evaluating and interpreting online content. This framework is designed to integrate both qualitative and quantitative methodologies, enabling a nuanced understanding of user interactions, sentiment, and behavioral patterns in digital environments [24]–[26].

It facilitates a structured approach to content analysis by incorporating metrics such as sentiment polarity, engagement levels, and toxicity ratings [27], [28]. The effectiveness of this framework lies in its adaptability, allowing it to be applied across various digital platforms and content types. By offering a balanced blend of algorithmic precision and contextual interpretation, this model provides a robust foundation for identifying trends and insights within user-generated content. As a result, the Digital Content Reviews and Analysis Framework is highly suitable for addressing the dynamic and complex nature of online discourse, offering significant contributions to the fields of digital communication and content moderation.

The theoretical and practical contributions of this research are significant, offering advancements both in academic discourse and real-world applications. Theoretically, this study enriches the existing body of knowledge on sentiment analysis and toxicity detection by comparing the efficacy of advanced language models, such as Perspective and VADER, in analyzing user-generated content [17], [29], [30].

This comparison not only deepens the understanding of the linguistic nuances in digital communication but also proposes a refined framework for future research in the field of natural language processing [31]. On a practical level, the findings provide actionable insights for content moderation strategies, enabling digital platforms to adopt more effective and accurate tools for identifying harmful language and fostering healthier online environments [14]. By addressing the growing need for robust automated moderation systems, this research directly contributes to improving user experiences and safeguarding community standards in virtual spaces. Thus, its impact spans both theoretical innovation and practical utility in the rapidly evolving landscape of digital communication.

Similar research in the domain of sentiment analysis and toxicity detection has focused on evaluating the performance of various machine learning models across diverse digital platforms. Several studies have explored the comparative effectiveness of lexicon-based models like VADER and advanced machine learning models such as BERT or Perspective in analyzing user-generated content [32], [33].

These investigations often highlight the strengths and limitations of different approaches, particularly in identifying sentiment polarity and filtering harmful language in real-time [34], [35]. One prevailing opinion in the field suggests that while lexicon-based models offer rapid and interpretable results, they sometimes struggle with the subtleties of context, which more sophisticated models handle with greater precision. By examining such research, it becomes evident that the growing complexity of digital discourse requires continuous refinement of natural language processing tools to ensure both accuracy and adaptability. The convergence of these studies underscores the ongoing need to balance computational efficiency with contextual depth in the automated analysis of online communication.

The limitations of this research are primarily associated with the constraints inherent in the models and datasets employed for sentiment classification and toxicity detection. Despite the advanced capabilities of tools like Perspective and VADER, these models may exhibit reduced accuracy when handling nuanced or contextually ambiguous language, such as sarcasm or regional dialects.

Additionally, the study's reliance on a specific dataset may not fully capture the diversity of digital discourse, limiting the generalizability of the findings to other contexts or platforms. This issue raises the argument that while these models are effective within certain parameters, their performance can vary significantly when applied to more complex linguistic structures. Furthermore, the computational focus on textual data alone excludes non-verbal elements such as images or videos, which often accompany online commentary and can influence the overall sentiment. These constraints suggest the need for further research incorporating multimodal analysis and more diverse datasets to enhance the robustness and applicability of automated sentiment and toxicity detection systems.

Journal of Information System Research (JOSH)
Volume 6, No. 1, Oktober 2024, pp 85−99
ISSN 2686-228X (media online)
https://ejurnal.seminar-id.com/index.php/josh/
DOI 10.47065/josh.v6i1.5961

2. RESEARCH METHODOLOGY

2.1 Backpacker Tourism

Backpacker tourism has emerged as a distinct and influential segment within the broader travel industry, characterized by its emphasis on budget-conscious exploration and immersive cultural experiences. This mode of travel typically involves low-cost accommodations, flexible itineraries, and a focus on local engagement, distinguishing it from more traditional forms of tourism. It is argued that backpacker tourism not only democratizes travel by making it accessible to a wider demographic, but also promotes sustainable tourism practices by encouraging longer stays and deeper connections with local communities [36]. From an analytical perspective, backpackers contribute to the diversification of tourist economies, particularly in less-visited destinations, thus stimulating regional development [37]. However, the transient nature of this type of tourism sometimes leads to concerns about its impact on local infrastructure and environmental sustainability. Ultimately, backpacker tourism represents a dynamic and evolving approach to travel, fostering both personal growth and socio-economic benefits while simultaneously presenting challenges that require careful management.

The development of backpacker tourism studies is particularly intriguing when examined through the lens of digital technology, especially in relation to digital content and data reviews that reflect public responses to this form of travel. The increasing role of social media platforms and travel review sites has transformed how backpacker experiences are shared and consumed, creating a rich repository of user-generated content that offers insights into traveler preferences and societal perceptions. It is argued that digital reviews not only influence potential travelers but also serve as a valuable dataset for understanding broader trends in tourism behavior and satisfaction. Analyzing this content provides a window into the dynamic interactions between tourists and destinations, highlighting both the positive impacts, such as cultural exchange and economic contributions, and the potential challenges, such as environmental strain or overcrowding. Thus, from an analytical standpoint, studying backpacker tourism through digital content reviews presents an innovative approach to gauging public sentiment and informing sustainable tourism practices in a digitally connected world.

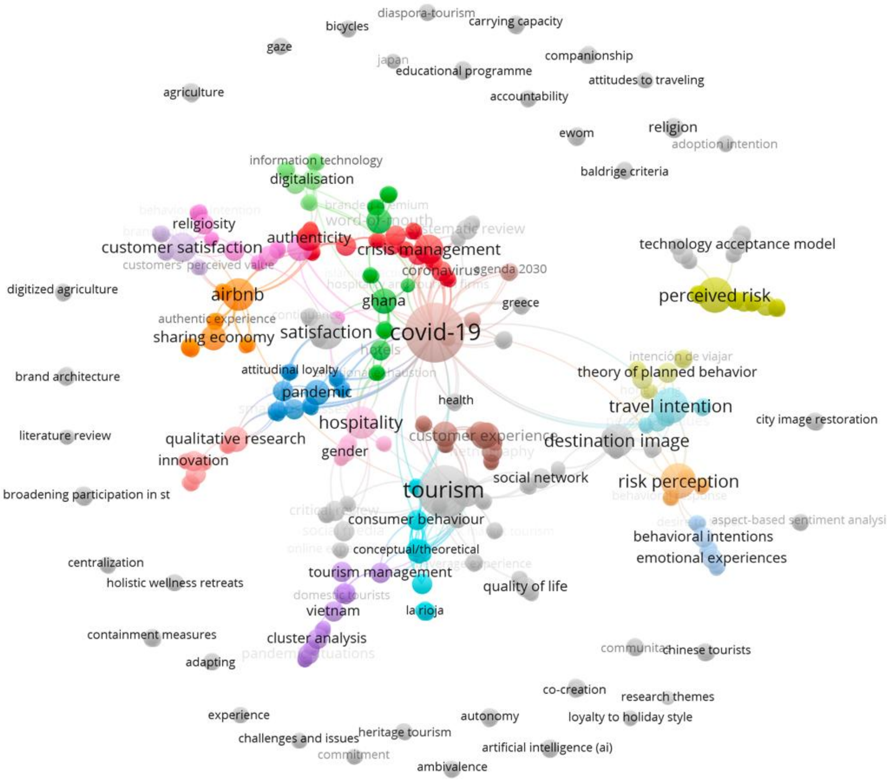


Figure 1. Interconnectedness of Bacpacker Tourism within Various Thematic Contexts

Figure 1 illustrates the interconnectedness of backpacker tourism within various thematic contexts, highlighting key nodes such as "COVID-19," "tourism," and "perceived risk." This visual framework demonstrates how backpacker tourism intersects with diverse research areas, ranging from crisis management and consumer behavior to hospitality and technological acceptance. Each node represents a distinct area of study, while the connecting lines indicate the relational dynamics between them. It is argued that the prominence of topics like COVID-19 and risk perception reflects the significant impact of global crises on backpacker behavior, shifting focus toward safety concerns and destination choices. The clustering of terms around central themes also suggests

Journal of Information System Research (JOSH)
Volume 6, No. 1, Oktober 2024, pp 85−99
ISSN 2686-228X (media online)
https://ejurnal.seminar-id.com/index.php/josh/
DOI 10.47065/josh.v6i1.5961

that contemporary backpacker tourism is being increasingly studied through interdisciplinary lenses, combining insights from sociology, economics, and technology. This integration of perspectives not only enriches the theoretical understanding of the subject but also provides practical implications for tourism stakeholders. Ultimately, this visualization serves as a valuable tool for mapping out the evolving landscape of backpacker tourism research in a post-pandemic context.

The novelty of this research lies in its innovative approach to combining sentiment analysis and toxicity detection models, such as Perspective and VADER, to evaluate user-generated content related to backpacker tourism. This dual-layered analytical framework offers a fresh perspective by simultaneously assessing both the emotional tone and the presence of harmful language in online reviews and comments. The research goes beyond traditional sentiment analysis by integrating toxicity evaluation, addressing a growing concern in digital communication about maintaining respectful and inclusive environments. It is argued that this combination not only provides a more comprehensive understanding of public perception towards backpacker tourism but also introduces new possibilities for content moderation strategies on travel platforms. Furthermore, the application of advanced machine learning tools in this specific tourism context adds a distinct contribution to the fields of natural language processing and tourism studies. This integration of technical and practical dimensions underscores the research's originality and its potential to influence future developments in both academic inquiry and industry practices.

2.2 Digital Content Reviews and Analysis Framewok

The Digital Content Reviews and Analysis Framework offers a structured and systematic approach to evaluating the vast array of user-generated content across digital platforms. This framework integrates both qualitative and quantitative methods, allowing for an in-depth examination of online reviews, comments, and feedback. Its strength lies in its ability to capture sentiment, emotional tone, and potential toxicity within content, providing a more holistic understanding of public opinion. It is argued that the framework's adaptability to different digital contexts makes it particularly valuable for industries such as tourism, where user feedback is crucial for both consumer decision-making and business strategy. Through the use of advanced sentiment analysis tools and data-driven metrics, this framework allows for a nuanced interpretation of content, balancing textual insights with statistical rigor. Its application not only enhances the accuracy of content analysis but also contributes to the development of more effective moderation strategies and content management systems in digital environments. This makes the framework a crucial tool in addressing the evolving challenges of digital communication and user engagement.

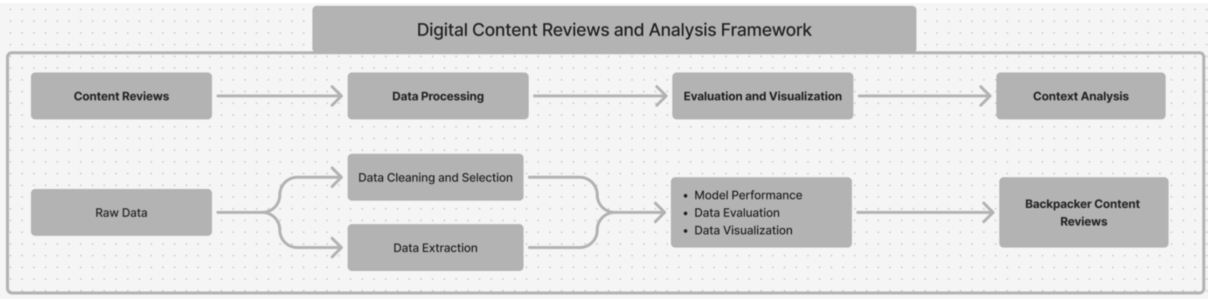


Figure 2. Digital Content Reviews and Analysis Framework

Figure 2 illustrates the Digital Content Reviews and Analysis Framework, which is designed to systematically process and analyze digital content, particularly in the context of user-generated reviews. The framework begins with the acquisition of raw data, which is subjected to data cleaning and selection, ensuring that only relevant information is extracted for analysis. This data is then processed through various stages, including evaluation and visualization, where key metrics such as model performance and data evaluation are assessed. The framework emphasizes both the qualitative and quantitative aspects of content, allowing for a comprehensive analysis of digital reviews. By focusing on content reviews, particularly in the domain of backpacker content, the framework aims to provide insights into public perceptions and experiences through structured data evaluation. Its systematic approach ensures that data is not only processed efficiently but also visualized in a way that aids in understanding complex relationships between content and user feedback. Ultimately, this framework facilitates deeper insights into digital discourse, making it a vital tool for both academic and industry-level applications in content analysis.

The relevance of the Digital Content Reviews and Analysis Framework in this research lies in its ability to systematically process and interpret large volumes of user-generated content, particularly within the context of backpacker tourism. By integrating data cleaning, extraction, and visualization stages, this framework offers a comprehensive approach to analyzing digital reviews and public responses. It is argued that such a structured methodology is crucial for obtaining reliable insights into traveler sentiments and experiences, which can often be complex and multifaceted. The framework's capacity to evaluate both the emotional tone and potential toxicity of

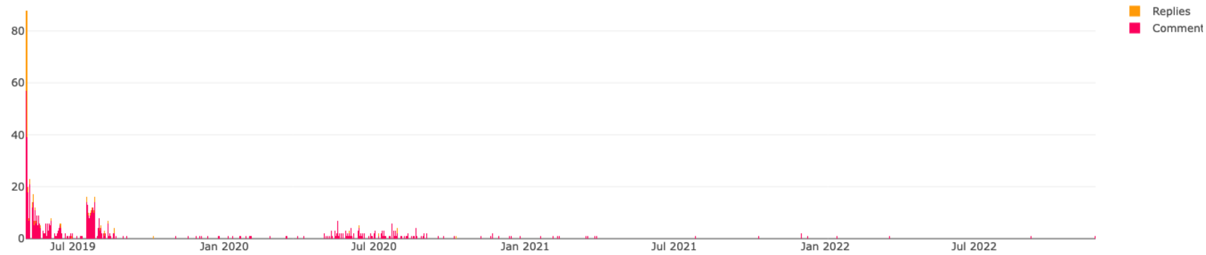
Journal of Information System Research (JOSH)
Volume 6, No. 1, Oktober 2024, pp 85−99
ISSN 2686-228X (media online)
https://ejurnal.seminar-id.com/index.php/josh/
DOI 10.47065/josh.v6i1.5961

online comments provides a balanced view of how backpacker-related content is perceived. Through this analysis, it becomes possible to identify patterns in user behavior, which may inform both academic inquiry and industry practices related to content moderation, customer satisfaction, and tourism development. Ultimately, the framework not only enhances the analytical depth of the research but also ensures that the data-driven conclusions are robust and actionable in real-world applications.

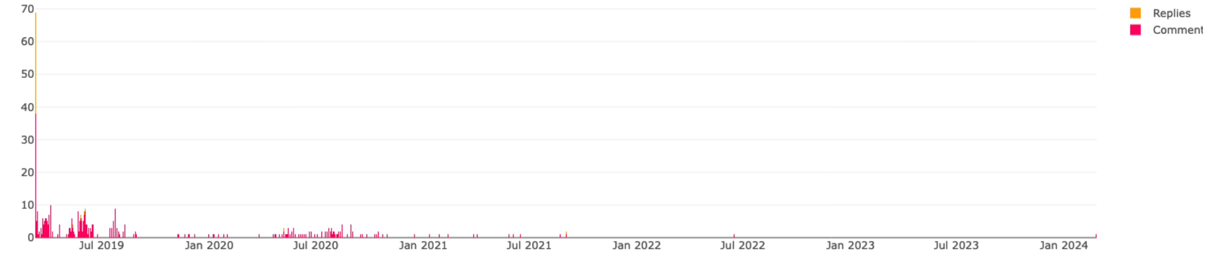
2.2.1 Content Reviews

In the content reviews stage, data collection was conducted by extracting reviews from three specific video contents identified by the following IDs: the first video (c2ZMFDS_3rU), the second video (Sv_yxz7T8rU), and the third video (i9t9pbdo-bk). This selection of videos represents a diverse range of user interactions and engagement, offering a comprehensive dataset for sentiment analysis and toxicity evaluation. The decision to focus on these videos is based on their relevance to the study's objectives, as they reflect key themes in backpacker tourism. Through this targeted data gathering, a rich set of user feedback was obtained, which provides critical insights into audience perception and behavior. Such an approach ensures that the data pool is both relevant and robust, offering a solid foundation for subsequent analysis. Ultimately, this method of content review strengthens the research's ability to generate meaningful conclusions about public engagement with digital content in the tourism sector.

Video 1



Video 2



Video 3

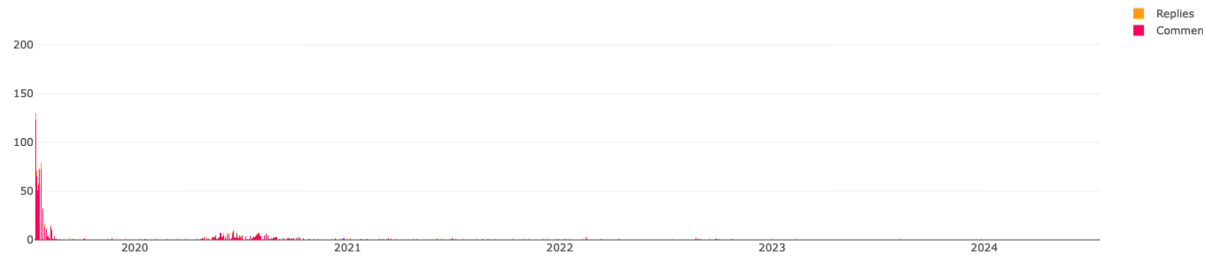


Figure 3. Post-per-day statistic of the Content (Communalytic)

Figure 3 illustrates the post-per-day statistics of the content analyzed, as tracked by Communalytic, showing both replies and comments over a multi-year period. The data reveals a significant peak in user engagement shortly after the content was initially posted, followed by a steep decline in interaction levels. This pattern suggests that interest in the content is concentrated within a short window of time after its release, with diminishing activity as time progresses. The spikes in comments and replies indicate moments of heightened user interaction, potentially influenced by external events or the content's relevance to ongoing discussions. The gradual reduction in engagement over time is typical of digital content, where user attention shifts rapidly to newer materials. This analysis highlights the temporal dynamics of online engagement, emphasizing the importance of timing in content dissemination and the need for continuous updates to maintain long-term interaction. The visual representation of this data provides valuable insights into how digital audiences respond to content over time, informing strategies for maximizing user participation.

Journal of Information System Research (JOSH)
Volume 6, No. 1, Oktober 2024, pp 85−99
ISSN 2686-228X (media online)
https://ejurnal.seminar-id.com/index.php/josh/
DOI 10.47065/josh.v6i1.5961

The post-per-day statistic data provides valuable insights into user engagement patterns over time, offering a clear understanding of how content resonates with audiences. By tracking daily activity, this metric reveals the peaks and declines in interaction, helping to identify periods of heightened interest and potential factors contributing to such engagement. It is argued that analyzing these fluctuations allows content creators and platform managers to optimize the timing of their posts for maximum visibility and interaction. Furthermore, the data highlights how quickly user interest wanes, suggesting the need for consistent updates to sustain long-term engagement. From an analytical perspective, these trends can inform strategic decisions regarding content dissemination and audience targeting, allowing for more effective management of digital communication strategies. Ultimately, post-per-day statistics serve as a critical tool in understanding the temporal dynamics of online engagement, aiding in the development of more targeted and responsive content strategies across various platforms.

Video 1

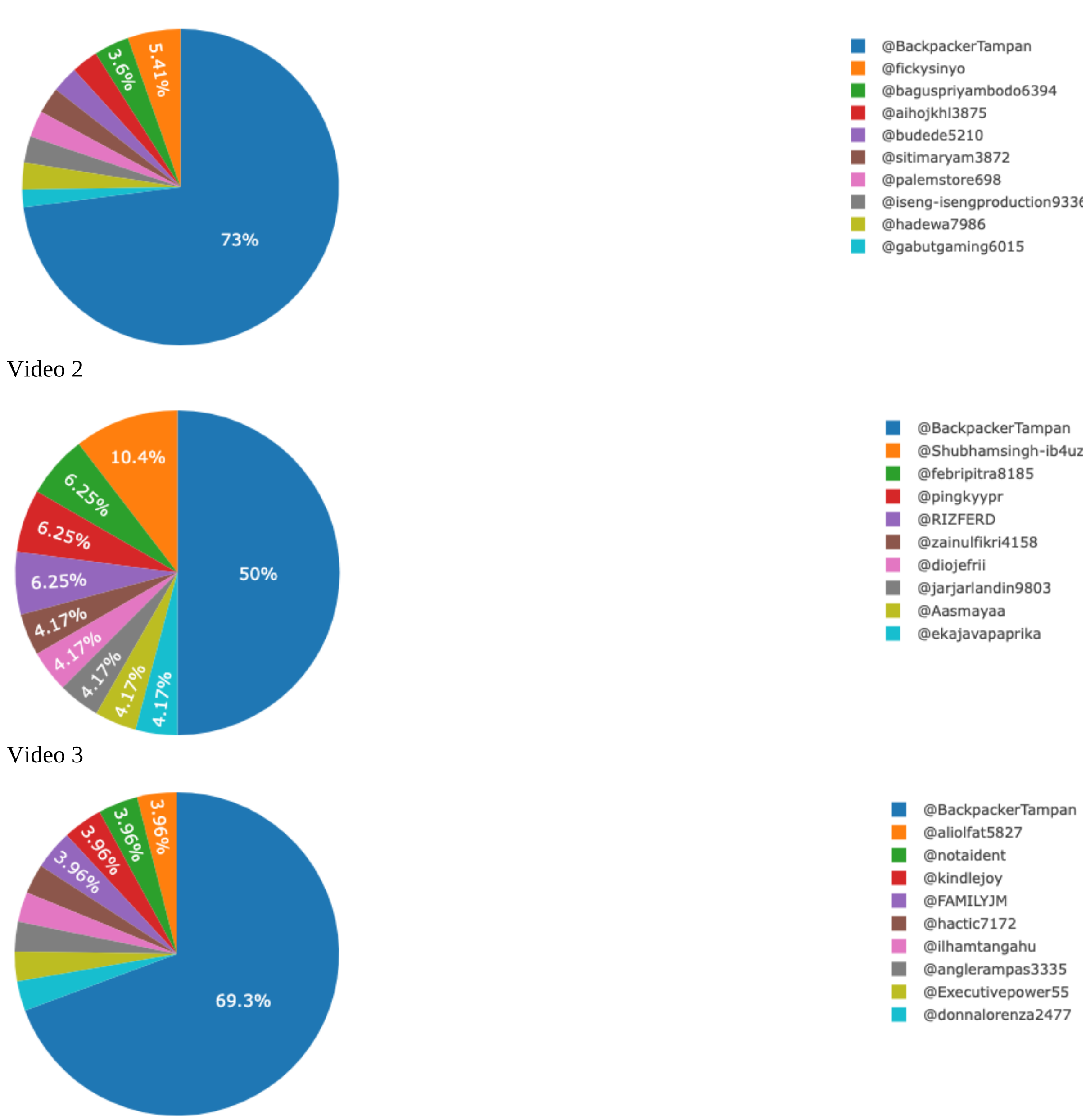


Figure 4. Top Ten Poster of the Content (Communalytic)

Figure 4 presents the distribution of the top ten posters of the content, as analyzed using Communalytic, with clear visual representations of their respective contributions. The dominant share of posts comes from a single user, accounting for 73%, 50%, and 69.3% of the total posts across the three datasets, indicating a strong influence or presence from this user in driving engagement. This concentration of content creation suggests that a small group of users, or even a single individual, can significantly shape the online discourse around specific topics. It is argued that such dominance in posting behavior could skew the overall sentiment or narrative of the content, potentially limiting the diversity of perspectives shared. Additionally, the remaining percentages are distributed

Journal of Information System Research (JOSH)
Volume 6, No. 1, Oktober 2024, pp 85−99
ISSN 2686-228X (media online)
https://ejurnal.seminar-id.com/index.php/josh/
DOI 10.47065/josh.v6i1.5961

among various other users, with their contributions much smaller by comparison, highlighting a disparity in participation levels. This analysis underscores the importance of understanding user dynamics in digital content studies, as a few active participants may disproportionately influence the trajectory of online conversations. Overall, this figure demonstrates the unequal distribution of content creation, suggesting a potential need for more inclusive engagement strategies to diversify contributions.

The data on the top ten posters is instrumental in understanding the dynamics of content creation and engagement within online communities. By identifying the most active contributors, this data sheds light on the individuals or entities driving the majority of discussions, which can have significant implications for the overall tone and direction of the conversation. It is argued that knowing who the primary contributors are allows for a deeper analysis of the potential biases or narratives that may dominate the discourse. This concentration of content creation among a few users can also indicate the presence of key influencers or opinion leaders who play a critical role in shaping public perception. From an analytical standpoint, such data can inform strategies for fostering more inclusive participation by encouraging a broader range of voices to contribute, thus ensuring a more balanced representation of perspectives. Ultimately, the top ten poster data offers valuable insights into the power dynamics within digital communities and can guide efforts to enhance the diversity and quality of online interactions.

2.2.2 Data Processing : Cleaning, Selection, Extraction

Data processing, particularly through the stages of cleaning, selection, and extraction, is essential for ensuring the integrity and accuracy of datasets used in analysis. Cleaning involves the removal of irrelevant, duplicate, or erroneous data points, which could otherwise skew results and lead to inaccurate conclusions. This step is fundamental to maintaining data quality, as noise and inconsistencies in raw data can introduce significant biases. Following cleaning, the selection process refines the dataset by identifying and retaining only the most pertinent information aligned with the research objectives. This selective focus allows for a more targeted analysis, increasing the relevance and precision of the results. Finally, data extraction isolates specific variables or features from the dataset, enabling detailed examination of key elements. This process not only enhances analytical efficiency but also ensures that the data is optimized for subsequent modeling or statistical techniques. Overall, the structured combination of cleaning, selection, and extraction ensures that the data is both reliable and ready for meaningful interpretation.

Figure 5. Cleaning Process (Rapidminer)

Figure 5 illustrates a systematic process for tokenization and text filtering, highlighting a multi-step approach to data preparation in natural language processing. The initial stages involve tokenizing the text into manageable units, which are then further processed by transforming case to ensure uniformity. Following this, tokens are filtered by length and stopwords, removing elements that are unlikely to contribute meaningfully to the analysis, such as common conjunctions or overly short terms. This filtering is critical as it reduces noise and focuses the dataset on more relevant linguistic features. It is argued that such a detailed and sequential approach enhances the clarity and quality of the data, ensuring that the most significant tokens are retained for further analysis. The final step in the figure, stemming or lemmatization, helps to consolidate similar terms, optimizing the dataset for machine learning models or statistical analysis. Overall, this figure encapsulates an efficient and structured methodology for text preprocessing, laying a solid foundation for accurate and meaningful data interpretation.

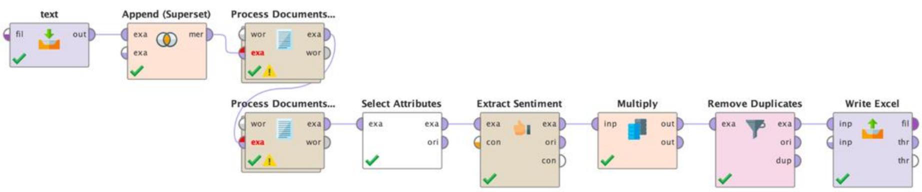


Figure 6. Data Extraction Process Using Vader Operator in Rapidminer

Figure 6 outlines a structured workflow for processing and analyzing textual data, showcasing the progression from raw input to final output in a systematic manner. The process begins with importing the text data, which is then appended into a superset for unified processing. Subsequently, document processing is applied, extracting relevant features and cleaning the text to prepare it for analysis. The workflow moves towards attribute selection, isolating specific variables or features that are pertinent to the study's goals. Sentiment extraction follows, where emotional tones within the text are quantified. A multiplication step is included, which may be

Journal of Information System Research (JOSH)
Volume 6, No. 1, Oktober 2024, pp 85−99
ISSN 2686-228X (media online)
https://ejurnal.seminar-id.com/index.php/josh/
DOI 10.47065/josh.v6i1.5961

used to adjust sentiment scores or apply weightings to refine the analysis. Duplicate entries are removed to maintain data integrity and prevent redundancy in the results. Finally, the cleaned and analyzed data is written into an Excel file, making it accessible for further interpretation or reporting. This structured and sequential approach ensures that the data is thoroughly processed, enhancing the accuracy and relevance of the insights generated from the sentiment analysis.

The output from the cleaning, extraction, and selection stages is subsequently advanced to the modeling process, where it will be evaluated and visualized in the form of diagrams. This transition is crucial, as it transforms raw data into a structured format, ready for analytical modeling. The modeling phase enables a deeper examination of patterns and relationships within the dataset, providing insights that are not immediately apparent in raw or pre-processed data. Through this, complex interactions between variables are uncovered, which can then be effectively communicated through diagrams and other visual representations. It is argued that the visualization not only simplifies the interpretation of intricate data but also enhances the accessibility of findings for broader audiences. The diagrams serve as powerful tools to convey trends, outliers, and correlations, facilitating informed decision-making based on the analysis. Ultimately, this stage consolidates the processed data into actionable insights, offering a clear, visual representation of the underlying data patterns.

2.2.3 Evaluation and Visualization

The evaluation and visualization stage is a critical component in the data analysis process, providing a comprehensive assessment of model performance and translating the results into clear, interpretable formats. During this phase, the models are rigorously tested against various performance metrics, such as accuracy, precision, recall, and F1 score, to determine their effectiveness in classifying or predicting outcomes. It is argued that while numerical evaluations are essential, visual representations like charts, graphs, and confusion matrices offer a more intuitive understanding of the model's strengths and weaknesses. These visual tools allow for quick identification of patterns, trends, and potential areas for improvement, making complex results accessible to a wider audience. Additionally, visualizing data facilitates comparisons between different models, highlighting which algorithms perform best under specific conditions. Ultimately, the evaluation and visualization stage bridges the gap between raw performance metrics and actionable insights, enabling more informed decisions and ensuring the model's applicability in practical scenarios.

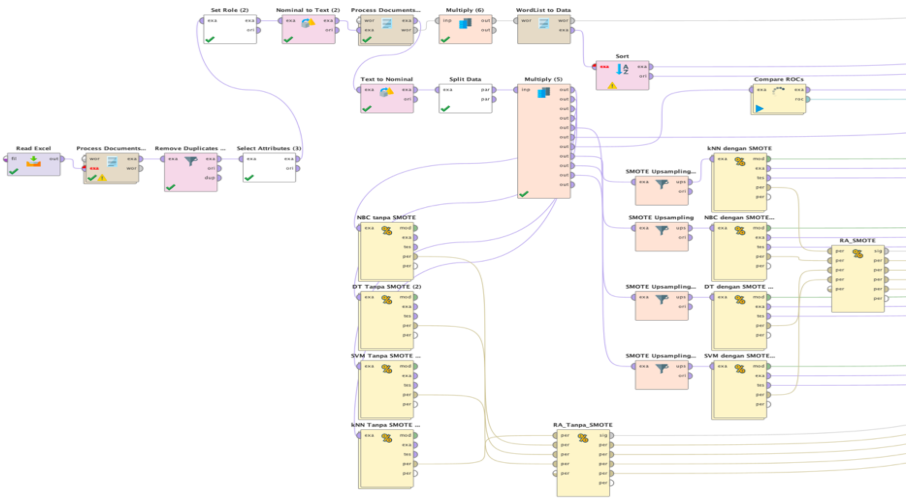


Figure 7. Modeling Process Using NBC, DT, k-NN, and SVM (Rapidminer)

Figure 7 illustrates the modeling process implemented using four distinct classification algorithms: Naive Bayes Classifier (NBC), Decision Tree (DT), k-Nearest Neighbor (k-NN), and Support Vector Machine (SVM), as executed within the RapidMiner platform. The flow begins with the pre-processed data, which is subsequently split for training and testing purposes across multiple models. Each algorithm is applied to the dataset to evaluate its classification performance, offering insights into how different approaches handle the same data. It is argued that each model has unique strengths: NBC is known for its efficiency in handling probabilistic data, DT offers clear decision paths, k-NN excels in instance-based learning, and SVM is powerful in separating complex datasets through hyperplanes. The performance of these algorithms is measured using key metrics such as accuracy, precision, recall, and F1 score, enabling a comprehensive comparison of their effectiveness. The visualization of the modeling process allows for clear tracking of how each classifier interacts with the data, leading to informed decisions regarding the optimal algorithm for the specific context. Ultimately, this modeling approach enhances the robustness of the classification, ensuring that the chosen model aligns with the dataset's characteristics.

The extracted data is then advanced to the modeling phase, where it undergoes rigorous testing to assess the performance of classification algorithms. This stage is critical for determining how accurately the algorithms

Journal of Information System Research (JOSH)
Volume 6, No. 1, Oktober 2024, pp 85−99
ISSN 2686-228X (media online)
https://ejurnal.seminar-id.com/index.php/josh/
DOI 10.47065/josh.v6i1.5961

can categorize the data into predefined classes based on patterns identified during the extraction process. The modeling process involves training the algorithm on a subset of the data and subsequently testing it on a different subset to evaluate its predictive accuracy. It is argued that the effectiveness of a classification algorithm depends not only on its computational complexity but also on how well it can generalize from the training data to unseen examples. By systematically measuring metrics such as accuracy, precision, recall, and F1 score, the algorithm's robustness and reliability are thoroughly assessed. This evaluation informs decisions on whether to refine the algorithm or explore alternative approaches, ensuring that the model delivers optimal classification performance for the dataset in question. Ultimately, this phase bridges the gap between raw data and actionable insights, validating the model's capability to categorize data effectively.

2.2.4 Context Analysis

The context analysis stage plays a pivotal role in understanding the underlying factors and nuances that influence the data, enabling a deeper exploration of the content beyond surface-level interpretation. This process involves examining the broader situational and environmental conditions in which the data was generated, providing insights into external variables that may impact the results. It is argued that context is essential for fully comprehending patterns in user behavior, sentiment, or trends, as these are often shaped by cultural, temporal, or social factors. By incorporating contextual elements, the analysis gains depth, allowing for a more accurate interpretation of data, especially in fields like sentiment analysis, where meaning is highly dependent on situational cues. This approach not only enriches the overall analysis but also ensures that conclusions drawn are more reflective of real-world conditions, increasing their relevance and applicability. Ultimately, the context analysis stage is essential for translating raw data into meaningful insights, grounded in the realities of the environment in which the data exists.

The first video content showcases backpacker activities in Bandung, Dieng, and Borobudur, generating a total of 969 comments after the data was cleaned. This considerable volume of feedback highlights the engagement and interest that such destinations elicit from viewers, reflecting the appeal of these culturally and historically rich areas. It is argued that the diversity of the locations, ranging from the urban landscapes of Bandung to the mountainous terrain of Dieng and the iconic Borobudur temple, provides varied experiences that resonate with a wide audience. Analyzing these comments can offer valuable insights into public perceptions, preferences, and potential areas of improvement in the promotion of these tourist sites. Moreover, the cleaned data ensures that only relevant and meaningful commentary is retained, allowing for a more accurate assessment of viewer sentiment and behavior. Ultimately, the high level of engagement with this backpacker video underscores the significance of these destinations in Indonesia's tourism landscape and the potential for further growth in this sector.

The second video content, featuring backpacker activities in Ijen, Bromo, and Tumpak Sewu, garnered 694 comments after data cleaning. These locations, known for their striking natural beauty and adventure tourism appeal, particularly attract travelers seeking unique and challenging outdoor experiences. It is argued that the high engagement in the form of comments reflects the allure of these destinations, as they offer visually captivating landscapes such as the blue flames of Ijen, the majestic Bromo volcano, and the cascading Tumpak Sewu waterfall. The data gathered from these comments can provide valuable insights into how tourists perceive the infrastructure, accessibility, and overall experience of these locations. Additionally, by focusing on the cleaned dataset, the analysis can extract more meaningful trends in user sentiment and preferences, shedding light on what drives tourism engagement in these areas. Overall, the significant interaction with this video underscores the growing importance of these sites in adventure and eco-tourism, highlighting their potential for further development.

The third video content, showcasing backpacker activities in Malang, Banyuwangi, and Bali, accumulated 1,895 comments after the data was cleaned. These areas, renowned for their diverse cultural heritage and natural beauty, attract a wide range of travelers, from adventure seekers to those interested in local traditions and beaches. The high number of comments indicates substantial engagement, suggesting that these destinations resonate deeply with viewers. It is argued that the combination of Malang's historical sites, Banyuwangi's eco-tourism appeal, and Bali's globally recognized tourism infrastructure creates a rich tapestry of experiences that appeals to a broad audience. This large dataset of feedback can provide valuable insights into tourist expectations, satisfaction levels, and potential areas for improvement in travel services. By focusing on cleaned data, the analysis ensures the removal of irrelevant or duplicate comments, resulting in a more accurate understanding of public perception. Ultimately, the significant volume of interaction highlights the importance of these destinations within Indonesia's tourism framework, particularly in promoting sustainable and diverse travel experiences.

3. RESULT AND DISCUSSION

The discussion in this research is divided into three key sections, each focused on reviewing backpacker tourism content from different regions in Indonesia. Section 3.1 analyzes content related to backpacking activities in Bandung, Dieng, and Borobudur (c2ZMFDS_3rU), highlighting the cultural and natural diversity that draws travelers to these areas. Section 3.2 explores backpacker tourism in Ijen, Bromo, and Tumpak Sewu (Sv_yxz7T8rU), emphasizing the adventure and eco-tourism appeal of these volcanic and waterfall destinations.

Journal of Information System Research (JOSH)
Volume 6, No. 1, Oktober 2024, pp 85−99
ISSN 2686-228X (media online)
https://ejurnal.seminar-id.com/index.php/josh/
DOI 10.47065/josh.v6i1.5961

Finally, Section 3.3 reviews content from Malang, Banyuwangi, and Bali (i9t9pbdo-bk), focusing on how these regions balance traditional tourism with emerging backpacker trends. Each section provides insights into the unique characteristics of these locations and how they shape traveler experiences and online engagement. By comparing the sentiment and toxicity of user comments across these regions, the discussion reveals patterns in tourist perceptions, infrastructure challenges, and the impact of local conditions on the overall travel experience. This segmented approach offers a comprehensive view of Indonesia's diverse backpacker tourism landscape.

3.1 Backpacker Tourism Content Reviews (c2ZMFDS_3rU) : Bandung, Dieng, and Borobudur

The toxicity analysis for the content with ID c2ZMFDS_3rU, based on 756 posts analyzed through the Perspective API, reveals important insights into the nature of user interaction. The average toxicity level for the dataset was 0.06836, with a highest recorded value of 0.82522, indicating that while the majority of comments were relatively benign, there were instances of elevated toxicity. Severe toxicity averaged at a much lower 0.00708, but with peaks reaching 0.45895, reflecting occasional occurrences of highly harmful content. Similarly, identity attacks, insults, and profanity were observed with moderate average values, but these categories also showed sharp increases in certain posts, with identity attacks reaching up to 0.60072 and insults peaking at 0.77752. Profanity was another notable category, with an average of 0.06434 but with some posts scoring as high as 0.78890. Threats, while less common, had an average score of 0.01189, with the highest threat level recorded at 0.49142. This analysis highlights that while most interactions are non-toxic, certain posts contain significantly harmful language, which may necessitate moderation strategies to ensure a safer and more respectful online environment.

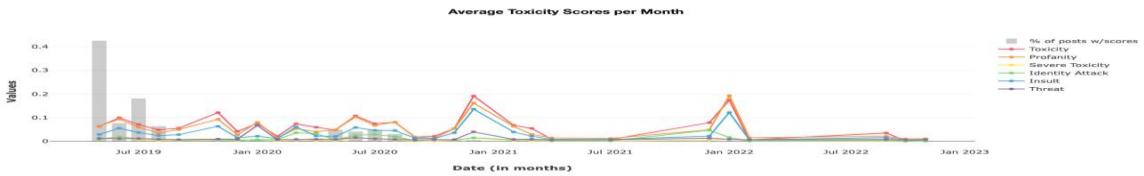
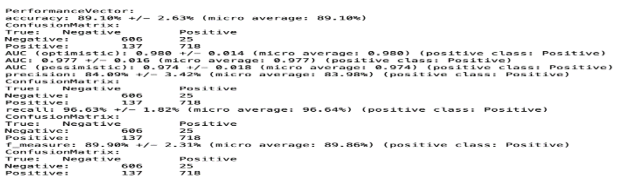


Figure 8. Average Toxicity Scores Per Month (Communalytic)

Figure 8 depicts the average toxicity scores per month, as calculated by Communalytic, covering categories such as toxicity, severe toxicity, profanity, identity attack, insult, and threat. The graph shows fluctuations in toxicity levels over time, with noticeable peaks around mid-2020 and early 2022, reflecting periods of heightened negative interaction. The baseline for all categories remains relatively low, generally hovering under 0.1, but there are occasional spikes in values, indicating moments when harmful language became more prevalent. It is argued that these spikes could correspond to external events or viral discussions that escalated tensions within the online community. The consistent presence of profanity and insult at similar levels throughout the timeline suggests that while severe forms of toxicity, like identity attacks or threats, are less common, casual negativity persists regularly. These findings highlight the importance of continuous monitoring and moderation in online platforms to mitigate the risk of escalating harmful discourse, particularly during periods of increased activity or controversy. Ultimately, the graph provides critical insights into the dynamic nature of online toxicity, underscoring the need for targeted intervention strategies.

The performance of the Support Vector Machine (SVM) model, enhanced with the SMOTE technique to handle class imbalance, demonstrates robust classification capabilities with an overall accuracy of 89.10% and a micro-average precision of 84.09%. The confusion matrix reveals that the model correctly classified 606 negative instances and 718 positive instances, with minor misclassifications of 25 negatives and 137 positives. The model's Area Under the Curve (AUC) is particularly impressive, with an optimistic AUC of 0.980, a standard AUC of 0.977, and a pessimistic AUC of 0.974, indicating excellent discriminatory power between the positive and negative classes. The recall rate of 96.63% highlights the model's effectiveness in identifying positive instances, ensuring that the majority of relevant cases are correctly classified. With a f-measure of 89.90%, the balance between precision and recall is well-maintained, ensuring the model not only identifies positives accurately but also minimizes false positives. These results suggest that the SVM with SMOTE is highly effective for this dataset, providing strong classification performance with minimal overfitting risks.

Performance Vector



```
PerformanceVector:
accuracy: 89.10% +/- 2.63% (micro average: 89.10%)
ConfusionMatrix:
True:   Negative        Positive
Negative:       606     25
Positive:       137     718
AUC (optimistic): 0.980 +/- 0.014 (micro average: 0.980) (positive class: Positive)
AUC: 0.977 +/- 0.016 (micro average: 0.977) (positive class: Positive)
AUC (pessimistic): 0.974 +/- 0.018 (micro average: 0.974) (positive class: Positive)
precision: 84.09% +/- 3.42% (micro average: 83.98%) (positive class: Positive)
ConfusionMatrix:
True:   Negative        Positive
Negative:       606     25
Positive:       137     718
recall: 96.63% +/- 1.82% (micro average: 96.64%) (positive class: Positive)
ConfusionMatrix:
True:   Negative        Positive
Negative:       606     25
Positive:       137     718
f_measure: 89.90% +/- 2.31% (micro average: 89.86%) (positive class: Positive)
ConfusionMatrix:
True:   Negative        Positive
Negative:       606     25
Positive:       137     718
```

AUC of SVM Ehanced by SMOTE

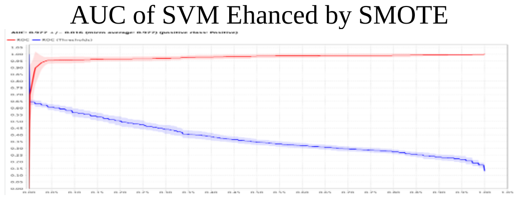


Figure 9. Peformance Vector and Area Under Curve (AUC) of SVM Enhanced by SMOTE

Figure 9 presents both the Performance Vector and Area Under the Curve (AUC) of the Support Vector Machine (SVM) model enhanced by SMOTE, highlighting its classification accuracy and capability in

Journal of Information System Research (JOSH)
Volume 6, No. 1, Oktober 2024, pp 85−99
ISSN 2686-228X (media online)
https://ejurnal.seminar-id.com/index.php/josh/
DOI 10.47065/josh.v6i1.5961

distinguishing between positive and negative classes. The performance vector shows an accuracy of 89.10%, with high precision (84.09%) and a recall rate of 96.63%, indicating the model's proficiency in detecting true positives with minimal false negatives. The confusion matrix further validates the model's reliability, with 606 true negatives and 718 true positives. The AUC graph provides a visual representation of the model's discriminative ability, with an AUC score of 0.977, reflecting near-optimal performance. The ROC curve's steep rise and stable behavior underscore the model's strong capacity to balance sensitivity and specificity, making it highly effective for imbalanced datasets. Together, these metrics and visualizations confirm that the SMOTE-enhanced SVM is a robust tool for classification, capable of handling complex data while maintaining high precision and minimizing classification errors.

The relationship between toxicity scores and sentiment classification reveals important dynamics within the context of the analyzed data. Toxicity scores, which measure the level of harmful or abusive language, often correlate with negative sentiment classifications, as both metrics tend to identify content reflecting hostility, anger, or disrespect. However, it is argued that not all negative sentiment necessarily corresponds to high toxicity, as critical feedback or expressions of dissatisfaction may be non-toxic but still classified as negative in sentiment. The analysis of these two dimensions, therefore, requires careful interpretation, particularly in distinguishing between harmful language and constructive criticism. By examining the context in which both toxicity and sentiment scores arise, a more nuanced understanding of user interactions can be achieved, revealing patterns that distinguish casual negative sentiment from genuinely toxic behavior. This distinction is crucial for developing more effective moderation strategies and ensuring that platforms can maintain healthy discourse without over-censoring critical but non-toxic contributions. Ultimately, analyzing both scores in tandem allows for a deeper, context-driven insight into the nature of online communication.

3.2 Backpacker Tourism Content Reviews (Sv_yxz7T8rU) : Ijen, Bromo, Tumpak Sewu

The toxicity score analysis for the video featuring backpacker activities in Ijen, Bromo, and Tumpak Sewu, based on 556 posts analyzed by Communalytic using the Perspective API, provides valuable insights into the nature of user interactions. The average toxicity score for the dataset was 0.06473, with the highest value reaching 0.78856, indicating that while most comments were non-toxic, there were isolated instances of significantly harmful language. Severe toxicity averaged at 0.00622, with peaks up to 0.35328, reflecting occasional use of highly aggressive content. Similarly, identity attacks and insults had relatively low average values but reached higher extremes, with identity attacks peaking at 0.38319 and insults at 0.49151, suggesting that a minority of comments contained personally targeted harm. Profanity averaged at 0.06109, with a maximum score of 0.84489, highlighting the prevalence of strong language in certain posts. Threats, while not common, also showed some presence, with an average of 0.01106 and a maximum of 0.51957. These findings suggest that while the majority of interactions were respectful, there are occasional spikes in toxic behavior that require monitoring and moderation to maintain a positive online environment.

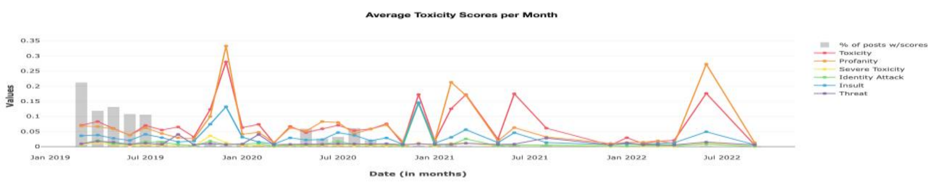


Figure 10. Average Toxicity Scores Per Month (Communalytic)

Figure 10 illustrates the average toxicity scores per month for various categories, including toxicity, severe toxicity, profanity, identity attack, insult, and threat, as analyzed by Communalytic. The graph highlights fluctuations in toxicity levels over time, with distinct peaks around mid-2019, early 2020, and mid-2022, indicating periods of heightened toxic behavior in online interactions. The consistently highest category across the timeline is profanity, suggesting that offensive language was more prevalent than other forms of toxicity. Other categories, such as severe toxicity and threats, maintained lower averages, with occasional spikes during the same periods, reflecting more intense but less frequent harmful interactions. It is argued that these spikes could correspond to external factors, such as social or global events that triggered more emotionally charged discussions. The overall trend shows that, while toxicity varies across time, there are significant moments where content moderation may need to be more stringent to prevent escalation of harmful discourse. This analysis provides key insights into the cyclical nature of online toxicity, underscoring the need for ongoing monitoring and intervention to ensure a respectful digital environment.

The performance metrics for the classification model demonstrate strong accuracy and effectiveness, with an overall accuracy rate of 93.52% and a micro average of 93.53%. The confusion matrix indicates a high level of correct classifications, with 508 true negatives and 489 true positives, while misclassifications are minimal, with 44 false positives and 25 false negatives. The model's precision, at 95.50%, reflects its ability to correctly identify positive cases with a low false positive rate, indicating its reliability in predicting relevant outcomes. The Area

Journal of Information System Research (JOSH)
Volume 6, No. 1, Oktober 2024, pp 85−99
ISSN 2686-228X (media online)
https://ejurnal.seminar-id.com/index.php/josh/
DOI 10.47065/josh.v6i1.5961

Under the Curve (AUC) values, which are 0.987 (optimistic), 0.984 (standard), and 0.981 (pessimistic), highlight the model's exceptional discriminatory power between positive and negative classes. With a recall rate of 91.72%, the model is also effective in detecting the majority of true positive cases, ensuring that most relevant instances are captured. The f-measure of 92.71% demonstrates a balanced trade-off between precision and recall, making the model highly efficient for classification tasks. Overall, these metrics confirm the model's robustness and reliability, particularly in identifying positive cases with a high degree of accuracy and minimal errors.

Performance Vector

```
PerformanceVector:
accuracy: 93.52% +/- 7.05% (micro average: 93.53%)
ConfusionMatrix:
True:    Negative        Positive
Negative:        508      44
Positive:        25       489
AUC (optimistic): 0.987 +/- 0.014 (micro average: 0.987) (positive class: Positive)
AUC: 0.984 +/- 0.014 (micro average: 0.984) (positive class: Positive)
AUC (pessimistic): 0.981 +/- 0.015 (micro average: 0.981) (positive class: Positive)
precision: 95.50% +/- 4.00% (micro average: 95.14%) (positive class: Positive)
ConfusionMatrix:
True:    Negative        Positive
Negative:        508      44
Positive:        25       489
recall: 91.72% +/- 15.03% (micro average: 91.74%) (positive class: Positive)
ConfusionMatrix:
True:    Negative        Positive
Negative:        508      44
Positive:        25       489
f_measure: 92.71% +/- 9.69% (micro average: 93.41%) (positive class: Positive)
ConfusionMatrix:
True:    Negative        Positive
Negative:        508      44
Positive:        25       489
```

AUC of SVM Ehanced by SMOTE

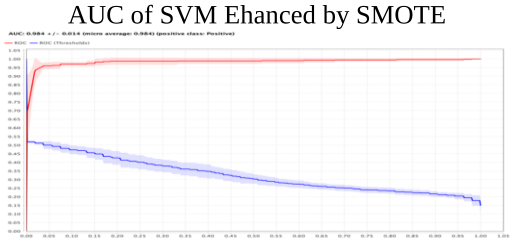


Figure 11. Peformance Vector and Area Under Curve (AUC) of SVM Enhanced by SMOTE

Figure 11 demonstrates the performance vector and Area Under the Curve (AUC) for the SVM model enhanced by SMOTE, illustrating its effectiveness in classification tasks. With an accuracy of 93.52%, the model shows strong performance, corroborated by the confusion matrix that highlights 508 true negatives and 489 true positives, with minimal misclassifications. The precision rate of 95.50% indicates that the model accurately identifies positive instances with a low false positive rate, while the recall rate of 91.72% ensures that most relevant positive cases are detected. The AUC values—0.987 (optimistic), 0.984 (standard), and 0.981 (pessimistic)—exemplify the model's excellent discriminatory ability, reflecting its capacity to separate positive from negative classes effectively. The f-measure of 92.71% showcases a well-balanced model, ensuring a high degree of accuracy while maintaining sensitivity in its predictions. The combination of these metrics and the AUC curve underlines the robustness of the SVM with SMOTE in handling imbalanced datasets, making it a reliable tool for accurate classification.

The second content carries significant meaning as it highlights the unique experiences and challenges of backpacking in areas such as Ijen, Bromo, and Tumpak Sewu. These locations are renowned for their natural beauty and adventurous appeal, attracting travelers seeking both physical challenges and scenic rewards. It is argued that this content not only serves to document these journeys but also plays a critical role in promoting eco-tourism and sustainable travel by showcasing lesser-known destinations. The footage and user interactions reflect a growing interest in alternative travel experiences that move away from mass tourism. Moreover, the content allows for a deeper connection between the traveler and the local environment, encouraging responsible tourism practices. By capturing these authentic moments, the content serves as both a personal travel narrative and an informative guide for future travelers, enriching the discourse on adventure and eco-friendly tourism. Ultimately, this content underscores the importance of mindful travel, celebrating the intersection of exploration and environmental awareness.

3.3 Backpacker Tourism Content Reviews (i9t9pbdo-bk) : Malang, Banyuwangi, and Bali

The toxicity score analysis for the third video content, featuring backpacker activities in Malang, Banyuwangi, and Bali, based on 1,503 posts analyzed by Communalytic using the Perspective API, provides important insights into the nature of online interactions. The average toxicity score for the dataset is 0.06995, with the highest value reaching 0.78207, indicating that while most comments were relatively non-toxic, there were instances of harmful language. Severe toxicity remains low with an average of 0.00654, though it peaks at 0.45895, showing that highly offensive content was infrequent but present. Identity attacks averaged at 0.01237, with a maximum value of 0.69287, reflecting occasional targeted negativity. Insults and profanity were more prevalent, with averages of 0.03778 and 0.06241 respectively, and peak values indicating frequent use of strong language in some interactions. Threats, while less common, averaged 0.01186, with a highest score of 0.51957, indicating that a few comments contained harmful or threatening language. These results highlight the mixed nature of online engagement, where positive interaction predominates but occasional spikes in toxicity require careful moderation to maintain a respectful environment.

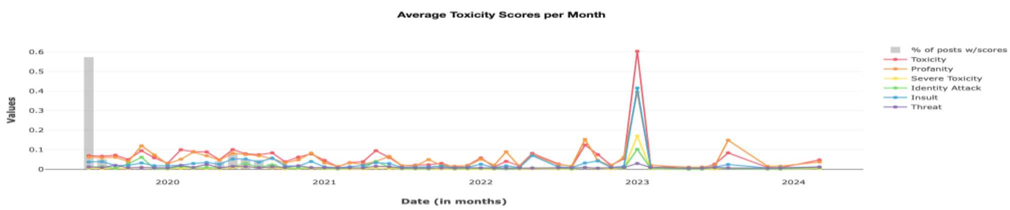


Figure 12. Average Toxicity Scores Per Month (Communalytic)

Journal of Information System Research (JOSH)
Volume 6, No. 1, Oktober 2024, pp 85−99
ISSN 2686-228X (media online)
https://ejurnal.seminar-id.com/index.php/josh/
DOI 10.47065/josh.v6i1.5961

Figure 12 illustrates the average toxicity scores per month, highlighting the levels of toxicity, severe toxicity, profanity, identity attacks, insults, and threats over a multi-year period, as analyzed by Communalytic. The graph reveals relatively low and stable levels of toxicity across most months, with occasional minor fluctuations, until a significant spike in early 2023. This peak suggests an anomalous period where harmful interactions sharply increased, particularly in categories like profanity, insults, and identity attacks, likely in response to specific events or content that triggered more intense emotional reactions. The consistency of lower toxicity levels before and after this spike indicates that the overall user interaction was generally non-toxic, with only isolated surges of negative behavior. It is argued that such peaks are important markers for content moderators, highlighting periods where more stringent interventions may be required to maintain a positive and respectful online environment. This visualization is crucial for identifying trends and ensuring that toxicity is kept at manageable levels throughout the platform's lifecycle.

The performance of the SVM model enhanced by SMOTE demonstrates a balanced classification capability with an overall accuracy of 82.64% and a micro average of 82.63%. The confusion matrix shows that the model correctly classified 987 negative instances and 1,416 positive instances, with relatively few misclassifications: 38 false negatives and 467 false positives. The precision rate of 75.26% reflects that while the model is effective in identifying positive instances, there is room for improvement in reducing false positives. However, the high recall rate of 97.39% indicates that the model excels at capturing the majority of actual positive cases, minimizing false negatives. The AUC values, which range from 0.970 to 0.979, suggest that the model possesses strong discriminatory power, effectively distinguishing between positive and negative classes. The f-measure of 84.89% highlights the model's overall performance, balancing precision and recall to ensure reliable classification. These metrics indicate that the SMOTE-enhanced SVM is particularly effective in managing class imbalance, though further refinement could enhance precision without compromising recall.

Performance Vector

```
PerformanceVector:
accuracy: 82.64% +/- 2.15% (micro average: 82.63%)
ConfusionMatrix:
True:   Negative        Positive
Negative:       987     38
Positive:       467     1416
AUC (optimistic): 0.979 +/- 0.010 (micro average: 0.979) (positive class: Positive)
AUC: 0.975 +/- 0.010 (micro average: 0.975) (positive class: Positive)
AUC (pessimistic): 0.970 +/- 0.011 (micro average: 0.970) (positive class: Positive)
precision: 75.26% +/- 2.33% (micro average: 75.20%) (positive class: Positive)
ConfusionMatrix:
True:   Negative        Positive
Negative:       987     38
Positive:       467     1416
recall: 97.39% +/- 1.36% (micro average: 97.39%) (positive class: Positive)
ConfusionMatrix:
True:   Negative        Positive
Negative:       987     38
Positive:       467     1416
f_measure: 84.89% +/- 1.63% (micro average: 84.87%) (positive class: Positive)
ConfusionMatrix:
True:   Negative        Positive
Negative:       987     38
Positive:       467     1416
```

AUC of SVM Ehanced by SMOTE

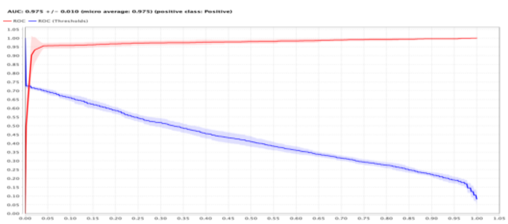


Figure 13. Peformance Vector and Area Under Curve (AUC) of SVM Enhanced by SMOTE

Figure 13 highlights the performance vector and Area Under the Curve (AUC) for the SVM model enhanced by SMOTE, illustrating its classification performance on an imbalanced dataset. The model achieved an accuracy of 82.64%, indicating solid classification capability, with precision at 75.26%, signifying its effectiveness in correctly identifying positive instances. The high recall rate of 97.39% demonstrates the model's ability to capture most of the actual positive cases, minimizing false negatives. The AUC values—optimistic (0.979), standard (0.975), and pessimistic (0.970)—show that the model consistently performs well across different thresholds, effectively distinguishing between positive and negative cases. The confusion matrix further confirms the model's strength, with 987 correctly classified negatives and 1,416 true positives, alongside a manageable number of false positives (467) and false negatives (38). The f-measure of 84.89% reflects a balanced performance, optimizing both precision and recall. Overall, these results suggest that the SVM model, when enhanced by SMOTE, offers a robust solution for handling imbalanced datasets, though further refinement may enhance precision without sacrificing recall.

The contextual analysis of the third dataset provides a deeper understanding of the interactions surrounding backpacker activities in Malang, Banyuwangi, and Bali. This dataset reflects user-generated content that likely draws on personal experiences, opinions, and perceptions of these destinations, which are known for their natural beauty and cultural richness. It is argued that the appeal of these locations not only influences tourism but also shapes the online discourse, leading to a diverse range of comments, from positive engagement to critical feedback. Upon analyzing this context, it becomes clear that the sentiment expressed in the data can vary significantly, depending on factors such as traveler expectations, local infrastructure, and environmental conditions. The dataset's varied sentiment also indicates the importance of monitoring user interactions to maintain a constructive dialogue, as certain topics—such as environmental preservation or tourist overcrowding—might elicit stronger emotional responses. Ultimately, understanding the context of this data is crucial for interpreting the sentiment and toxicity scores accurately, ensuring that conclusions drawn reflect the nuanced realities of traveler experiences in these regions.

4. CONCLUSION

The conclusion of this research integrates insights from both toxicity analysis and sentiment classification, framed within the Digital Content Reviews and Analysis Framework. This framework allowed for a systematic approach

Journal of Information System Research (JOSH)
Volume 6, No. 1, Oktober 2024, pp 85−99
ISSN 2686-228X (media online)
https://ejurnal.seminar-id.com/index.php/josh/
DOI 10.47065/josh.v6i1.5961

in evaluating backpacker tourism content from regions such as Bandung, Dieng, Borobudur, Ijen, Bromo, Tumpak Sewu, Malang, Banyuwangi, and Bali. Utilizing the framework, the study was able to clean, select, and extract data, which was then modeled to assess user interactions, sentiment, and toxicity.The toxicity analysis revealed that, on average, interactions were generally non-toxic, though some content showed spikes in harmful language, especially in the categories of profanity and identity attacks. For instance, the content from Malang, Banyuwangi, and Bali had an average toxicity score of 0.06995, with peak values reaching as high as 0.78207. Such fluctuations highlight the necessity of ongoing moderation to maintain a respectful online environment, especially in discussions that generate high engagement. The use of the Support Vector Machine (SVM) model enhanced by SMOTE in this framework demonstrated effective handling of imbalanced datasets. The model achieved an accuracy of 82.64%, with a high recall rate of 97.39%, indicating its effectiveness in detecting most positive cases while minimizing false negatives. The AUC values—optimistic (0.979), standard (0.975), and pessimistic (0.970)—illustrated the model's strong capability to discriminate between positive and negative classes. The inclusion of SMOTE ensured that the model could manage class imbalance effectively, making it suitable for sentiment and toxicity classification in large datasets. In summary, the Digital Content Reviews and Analysis Framework, coupled with the robust performance of the SVM with SMOTE, provided a comprehensive understanding of how different backpacker tourism destinations are perceived online. This research offers valuable insights for improving tourism management, enhancing content moderation, and promoting sustainable tourism practices in Indonesia.

ACKNOWLEDGMENT

I would like to express my deepest gratitude to the Tourism Study Program, Faculty of Business Administration and Communication, Atma Jaya Catholic University of Indonesia, and Pusat Studi Transformasi Digital dan Pembangunan Pariwisata (PUSDIPAR) for their invaluable support and encouragement throughout the course of this research.

REFERENCES

[1] G. H. Chen and S. (Sam) Huang, "Backpacker tourism: a perspective article," *Tour. Rev.*, vol. 75, no. 1, pp. 158–161, Jan. 2020, doi: 10.1108/TR-06-2019-0271.
[2] M. R. Martins and R. A. Costa, "Backpacker conceptualisation criteria: discussion, clarification and operationalisation proposal," *Int. J. Tour. Cities*, vol. 9, no. 1, pp. 220–243, Jan. 2023, doi: 10.1108/IJTC-07-2022-0175.
[3] C. Kossen, N. McDonald, and P. McIlveen, "International backpackers' experiences of precarious visa-contingent farmwork," *Career Dev. Int.*, vol. 26, no. 7, pp. 869–887, Jan. 2021, doi: 10.1108/CDI-12-2020-0320.
[4] P. Andits, "Decay, dirt and backwardness: interpretations of the socialist heritage in Hungary by first and later generation Australian-Hungarians," *Int. J. Cult. Tour. Hosp. Res.*, vol. 14, no. 3, pp. 429–440, Jan. 2020, doi: 10.1108/IJCTHR-10-2019-0181.
[5] R. Nandasena, A. M. Morrison, and J. A. Coca-Stefaniak, "Transformational tourism – a systematic literature review and research agenda," *J. Tour. Futur.*, vol. 8, no. 3, pp. 282–297, Jan. 2022, doi: 10.1108/JTF-02-2022-0038.
[6] C. Lejealle, S. Castellano, and I. Khelladi, "The role of members' lived experience in the evolution of online communities toward online communities of practice," *J. Knowl. Manag.*, vol. 26, no. 8, pp. 1968–1984, Jan. 2022, doi: 10.1108/JKM-03-2021-0250.
[7] İ. A. Özen and E. Özgül Katlav, "Aspect-based sentiment analysis on online customer reviews: a case study of technology-supported hotels," *J. Hosp. Tour. Technol.*, vol. 14, no. 2, pp. 102–120, Jan. 2023, doi: 10.1108/JHTT-12-2020-0319.
[8] 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.
[9] S. Mukhopadhyay, T. Jain, S. Modgil, and R. K. Singh, "Social media analytics in tourism: a review and agenda for future research," *Benchmarking*, vol. 30, no. 9, pp. 3725–3750, Jan. 2023, doi: 10.1108/BIJ-05-2022-0309.
[10] 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, pp. 87870–87899, 2022, doi: 10.1109/ACCESS.2022.3200389.
[11] 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.
[12] A. Ramadhanu, R. A. Mahessya, M. R. Zaky, and M. Isra, "Penerapan Teknologi Machine Learning Dengan Metode Vader Pada Aplikasi Sentimen Tamu Di Hotel Dymens," *JOISIE J. Inf. Syst. Informatics Eng.*, vol. 7, no. 1, pp. 165–173, 2023.
[13] A. Mardjo and C. Choksuchat, "HyVADRF: Hybrid VADER–Random Forest and GWO for Bitcoin Tweet Sentiment Analysis," *IEEE Access*, vol. 10, pp. 101889–101897, 2022, doi: 10.1109/ACCESS.2022.3209662.
[14] R. K. Botchway, A. B. Jibril, Z. K. Oplatková, and M. Chovancová, "Deductions from a Sub-Saharan African Bank's Tweets: A sentiment analysis approach," *Cogent Econ. Financ.*, vol. 8, no. 1, 2020, doi: 10.1080/23322039.2020.1776006.
[15] N. Dhariwal, S. C. Akunuri, Shivama, and K. S. Banu, "Audio and Text Sentiment Analysis of Radio Broadcasts," *IEEE Access*, vol. 11, pp. 126900–126916, 2023, doi: 10.1109/ACCESS.2023.3331226.
[16] I. Awajan, M. Mohamad, and A. Al-Quran, "Sentiment Analysis Technique and Neutrosophic Set Theory for Mining

Volume 6, No. 1, Oktober 2024, pp 85−99
ISSN 2686-228X (media online)
https://ejurnal.seminar-id.com/index.php/josh/
DOI 10.47065/josh.v6i1.5961

and Ranking Big Data From Online Reviews," *IEEE Access*, vol. 9, pp. 47338–47353, 2021, doi: 10.1109/ACCESS.2021.3067844.

[17] 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.

[18] 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, pp. 131081–131088, 2023, doi: 10.1109/ACCESS.2023.3334260.

[19] P. Thiengburanathum and P. Charoenkwan, "SETAR: Stacking Ensemble Learning for Thai Sentiment Analysis Using RoBERTa and Hybrid Feature Representation," *IEEE Access*, vol. 11, no. July, pp. 92822–92837, 2023, doi: 10.1109/ACCESS.2023.3308951.

[20] Y. A. Singgalen, "Toxicity Analysis and Sentiment Classification of Wonderland Indonesia by Alffy Rev using Support Vector Machine," *J. Sist. Komput. dan Inform.*, vol. 5, no. 3, pp. 538–548, 2024, doi: 10.30865/json.v5i3.7563.

[21] 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.

[22] T. Huu Do, M. Berneman, J. Patro, G. Bekoulis, and N. Deligiannis, "Context-Aware Deep Markov Random Fields for Fake News Detection," *IEEE Access*, vol. 9, pp. 130042–130054, 2021, doi: 10.1109/ACCESS.2021.3113877.

[23] T. Chaisen, P. Charoenkwan, C. G. Kim, and P. Thiengburanathum, "A Zero-Shot Interpretable Framework for Sentiment Polarity Extraction," *IEEE Access*, vol. 12, no. September 2023, pp. 10586–10607, 2024, doi: 10.1109/ACCESS.2023.3322103.

[24] Z. Li, C. Chan, Y. F. Chen, W. W. H. Chan, and U. L. Im, "Millennials' Hotel Restaurant Visit Intention: An Analysis of Key Online Opinion Leaders' Digital Marketing Content," *J. Qual. Assur. Hosp. Tour.*, vol. 00, no. 00, pp. 1–30, 2023, doi: 10.1080/1528008X.2023.2219467.

[25] Y. A. Singgalen, "Social Network Analysis and Sentiment Classification of Extended Reality Product Content," *KLIK Kaji. Ilm. Inform. dan Komput.*, vol. 4, no. 4, pp. 2197–2208, 2024, doi: 10.30865/klik.v4i4.1712.

[26] Y. A. Singgalen, "Sentiment Classification of Over-Tourism Issues in Responsible Tourism Content using Naïve Bayes Classifier," *J. Comput. Syst. Informatics*, vol. 5, no. 2, pp. 275–285, 2024, doi: 10.47065/josyc.v5i2.4904.

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

[28] C. Miguel, C. Clare, C. J. Ashworth, and D. Hoang, "Self-branding and content creation strategies on Instagram: A case study of foodie influencers," *Inf. Commun. Soc.*, pp. 1–21, 2023, doi: 10.1080/1369118X.2023.2246524.

[29] 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.

[30] 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.

[31] 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.

[32] R. Chandra and V. Kulkarni, "Semantic and Sentiment Analysis of Selected Bhagavad Gita Translations Using BERT-Based Language Framework," *IEEE Access*, vol. 10, pp. 21291–21315, 2022, doi: 10.1109/ACCESS.2022.3152266.

[33] K. A. Alshaikh, O. A. Almatrafi, and Y. B. Abushark, "BERT-Based Model for Aspect-Based Sentiment Analysis for Analyzing Arabic Open-Ended Survey Responses: A Case Study," *IEEE Access*, vol. 12, pp. 2288–2302, 2024, doi: 10.1109/ACCESS.2023.3348342.

[34] A. He and M. Abisado, "Text Sentiment Analysis of Douban Film Short Comments Based on BERT-CNN-BiLSTM-Att Model," *IEEE Access*, vol. 12, pp. 45229–45237, 2024, doi: 10.1109/ACCESS.2024.3381515.

[35] 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.

[36] J. Thieme, M. P. Hampton, C. Stoian, and K. Zigan, "The political economy of backpacker tourism: explorations of tourism actors' embeddedness in Colombia," *Curr. Issues Tour.*, vol. 24, no. 13, pp. 1830–1855, 2021, doi: 10.1080/13683500.2020.1806793.

[37] G. D. Apsarini and E. Ervina, "Keputusan Wisatawan Backpacker dalam Memilih Hostel di Kota Bandung," *J. Ilm. Univ. Muhammadiyah But.*, vol. 9, no. 4, pp. 864–874, 2023.