Implementation of Global Vectors for Word Representation (GloVe) Model and Social Network Analysis through Wonderland Indonesia Content Reviews

Yerik Afrianto Singgallen
Faculty of Business Administration and Communication, Tourism Study Program, Atma Jaya Catholic University of Indonesia, Jakarta, Indonesia
Email: yerik.afrianto@atmajaya.ac.id
Correspondence Author Email: yerik.afrianto@atmajaya.ac.id
Submitted: 20/03/2024; Accepted: 31/03/2024; Published: 31/03/2024

Abstract—Integrating the Global Vectors for Word Representation (GloVe) Model with Social Network Analysis presents a promising approach for extracting nuanced semantic relationships from Wonderland Indonesia's content reviews. However, the lack of comprehensive studies exploring the effectiveness of this integration, specifically within the context of Wonderland Indonesia's content reviews, necessitates focused research to uncover its potential impact and applications. This study investigates the reception and impact of the “Wonderland Indonesia” video content by Alffy Rev ft. Novia Bachmid (Chapter 1) within the YouTube community using a comprehensive methodology based on CRoss-Industry Standard Process for Data Mining (CRISP-DM), topic analysis, and Social Network Analysis (SNA). Through topic analysis, the content's main themes and narrative elements were identified, shedding light on its storytelling effectiveness. Furthermore, sentiment analysis using Vader was conducted on 2204 out of 24185 posts, revealing that 1369 (92%) exhibited positive sentiment, 427 (31.19%) had neutral sentiment, and 850 (62.09%) contained negative sentiment. Additionally, sentiment analysis using TextBlob was performed on the same subset of posts, with 1369 (40) posts exhibiting positive sentiment, 599 (43.75%) with neutral sentiment, and 730 (53.32%) expressing negative sentiment. Notably, metrics such as toxicity (highest value: 0.90780) and severe toxicity (highest value: 0.95021) exhibited varying prominence within the analyzed content. These findings enable targeted interventions and content moderation strategies to promote healthier online discourse. The SNA uncovered intricate social dynamics and interaction patterns among viewers, emphasizing the video's ability to foster engagement and community interaction. This study underscores the significance of creative storytelling and community engagement strategies in digital content creation, with implications for audience participation and community development within the digital sphere. Future research could explore the longitudinal effects of such content strategies on audience retention and community engagement.

Keywords: Wonderland Indonesia; Alffy Rev; Creative Content; Social Network Analysis; Topic Analysis

1. INTRODUCTION
Implementing the Global Vectors for Word Representation (GloVe) model can enhance understanding and analysis within diverse linguistic contexts. Wonderland Indonesia Music-Video Content Reviews serve as a rich dataset for exploring the application of GloVe in capturing semantic relationships and sentiments embedded in user-generated content. However, while GloVe excels in capturing global word co-occurrence statistics, its efficacy within the specific domain of social network analysis in the context of Indonesian music video reviews remains largely unexplored. Thus, investigating the integration of GloVe with social network analysis techniques offers a promising avenue for uncovering more profound insights into the dynamics of online interactions and content perceptions within the Wonderland Indonesia music video community.

The challenge for content creators lies in delivering educational content using storytelling techniques through audio-visual media to engage and captivate audiences effectively. Incorporating educational material into narratives enhances comprehension and retention by tapping into emotional and cognitive processes [1]. Furthermore, leveraging audio-visual mediums allows for multi-sensory experiences that deepen understanding and facilitate immersive learning environments [2]. Thus, by skillfully blending storytelling elements with educational content through audio-visual media, content creators effectively address the complexities of conveying information while fostering audience engagement and learning [3], [4].

Creative content has the potential to actively engage with the community as a targeted audience, fostering meaningful interactions and building rapport. Through innovative storytelling techniques and compelling visuals, content creators capture the attention of their intended audience, sparking interest and prompting participation [5]. Moreover, by addressing community interests and concerns, creative content establishes a sense of belonging and relevance, strengthening community ties and enhancing user engagement [6]. Consequently, leveraging creativity in content creation is a powerful tool for community engagement, facilitating meaningful connections and dialogue between content creators and their audiences.

One of the creative content pieces in the audio-visual medium, disseminated through the Alffy Rev channel, titled "Wonderland Indonesia," has garnered significant attention from YouTube users, amassing an impressive 56,646,952 views as of August 17, 2021, accompanied by 246,719 comments. This remarkable viewership and engagement highlight the effectiveness of the content in captivating audiences and stimulating discourse within the YouTube community. The success of "Wonderland Indonesia" underscores the potential of creative storytelling and visual presentation to resonate with audiences, demonstrating the power of compelling content to attract widespread attention and foster meaningful interactions [7], [8]. The urgency of this research lies in...
comprehending the behavior of YouTube users as a representation of the public in critiquing both creative and educational national content through videos. Understanding user engagement, preferences, and reactions to such content provides invaluable insights into audience reception and the effectiveness of content dissemination strategies [9]–[11]. Moreover, delving into user behavior sheds light on societal attitudes, preferences, and cultural dynamics, which are crucial for tailoring content to resonate with target audiences effectively [12]–[14]. Thus, by investigating YouTube user behavior in critiquing creative and educational content, this research optimizes content creation and fosters meaningful interactions between creators and their audiences [15].

The method offered is CRISP-DM (Cross-Industry Standard Process for Data Mining), facilitated through social network and topic analysis. CRISP-DM offers a comprehensive framework for understanding, preparation, modeling, and evaluation [16]–[18]. By incorporating social network analysis, this research uncovers the intricate relationships and interactions within online communities, offering insights into the dissemination and reception of content [19]–[22]. Additionally, employing topic analysis allows for the systematic examination of prevalent themes and discourse patterns, aiding in interpreting user-generated content [23]–[26]. Thus, integrating CRISP-DM with social network and topic analysis methodologies offers a comprehensive approach to understanding and interpreting user behavior and content dynamics in online platforms. Analyzing social networks and topic patterns is essential for gaining insights into the intricate web of content reviewers within video content. Understanding the dynamics of social networks aids in uncovering the interconnectedness and influence distribution among individuals engaging in content critique [27]–[31]. Moreover, topic analysis provides a structured approach to discerning prevalent themes and discourse trends within these networks.

This research unravels the underlying content evaluation and perception patterns by employing advanced social network and topic analysis methodologies, facilitating informed decision-making processes in content creation and dissemination strategies [32]–[34]. In conclusion, integrating social network and topic analysis offers a comprehensive framework for elucidating the intricate mechanisms governing content evaluation within online video platforms.

Theoretical and practical implications of Social Network Analysis (SNA) and Topic Analysis utilizing the CRISP-DM methodology are profound in contemporary research endeavors. By employing SNA, this research delves into online communities' intricate structures and dynamics, shedding light on information dissemination patterns, influence networks, and community engagement. Concurrently, Topic Analysis facilitates systematically exploring thematic content within these networks, enabling the discernment of prevailing discourses and user interests. Through the structured framework of CRISP-DM, these methodologies are integrated seamlessly, offering a comprehensive approach to data analysis and interpretation. Consequently, applying SNA and Topic Analysis within the CRISP-DM framework advances theoretical understandings of online behavior. It provides actionable insights for practitioners and policymakers in various domains, from digital marketing to public health interventions. This research's limitation is rooted in applying the CRISP-DM methodology through Social Network Analysis (SNA) and topic analysis. While these methodologies offer valuable insights into online behavior and content dynamics, their efficacy may be constrained by certain factors such as data availability, sample representativeness, and algorithmic biases. Additionally, the scope of the discussion is confined to analyzing a specific video identified by its unique identifier, aKtb7Y3qOck. Despite these limitations, rigorous methodological approaches and targeted analyses yield valuable findings and advance scholarly discourse in digital media research.

2. RESEARCH METHODOLOGY

2.1 Research Gap Analysis using Vosviewer

In this research endeavor, gap analysis is imperative to elucidate the dynamic interplay between content analysis, social media, and digital platforms. Through a systematic examination of existing literature and empirical data, gap analysis is a pivotal tool to uncover the prevailing trends and intricate relationships within these domains. This research discerns unexplored avenues and potential research trajectories by meticulously identifying disparities and inadequacies. Consequently, this methodological approach fosters a comprehensive understanding of the evolving landscape of digital communication, facilitating informed decision-making and scholarly advancement.

![Figure 1. Network and Density Visualization using Vosviewer](image-url)

Copyright © 2024 Yerik Afrianto Singgelen, Page 560
This Journal is licensed under a Creative Commons Attribution 4.0 International License
Figure 1 shows the network and density visualization using Vosviewer. The intricate interconnections between social media dynamics, content analysis methodologies, and the evolving landscape of YouTube and creative industries become discernible through a meticulous examination of gap analysis outcomes. This research gains valuable insights into how content creation, dissemination, and audience engagement intersect across digital platforms by scrutinizing these interrelationships. This analysis underscores the symbiotic nature of social media ecosystems, content production strategies, and the burgeoning influence of platforms like YouTube within the creative economy. Consequently, leveraging the insights derived from gap analysis allows for a nuanced understanding of the evolving dynamics within digital realms, paving the way for informed decision-making and strategic interventions to navigate this complex terrain effectively.

Considering the burgeoning research landscape concerning digital content, which underscores the intricate interplay among creative industries, content analysis methodologies, social media platforms, and YouTube, this study proposes the exploration of network dynamics through Social Network Analysis (SNA) alongside sentiment-based topic analysis using the CRISP-DM framework. By integrating these analytical approaches, this research unravels the complex web of interactions and sentiments within digital ecosystems, shedding light on emergent patterns and trends. This methodological synthesis offers a comprehensive lens to dissect the multifaceted dynamics of contemporary digital landscapes, thereby enriching scholarly discourse and informing strategic interventions for stakeholders across various industries.

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

The framework employed for identifying social network users and topic modeling is the Cross-Industry Standard Process for Data Mining (CRISP-DM). By adopting CRISP-DM, this research systematically navigates the various stages of data exploration, preparation, modeling, evaluation, and deployment. This structured approach ensures a methodical and efficient analysis of social network data, enabling uncovering underlying patterns, relationships, and insights. Leveraging the robust methodology of CRISP-DM enhances the rigor and reliability of the research process and fosters reproducibility and scalability in addressing complex research questions.

Figure 2. Implementation of CRISP-DM in Sentiment Classification

Figure 2 shows the CRISP-DM framework. The excellence of the Cross-Industry Standard Process for Data Mining (CRISP-DM) in Social Network Analysis (SNA) and Topic Analysis lies in its structured methodology, which offers a systematic approach to data exploration, preparation, modeling, evaluation, and deployment. This structured framework ensures a methodical examination of complex datasets, allowing the uncovering of intricate patterns and relationships within social networks while simultaneously facilitating the identification and analysis of topics and sentiments. By embracing CRISP-DM, this research effectively navigates the intricacies of data mining processes, enhancing the rigor and reliability of their analyses and ultimately contributing to more informed decision-making and strategic interventions.
2.2.1 Business Understanding

At the business understanding stage, an investigation into Social Network Analysis (SNA) and Topic Analysis is conducted on the video content from the creative industry of Aliffy Rev, explicitly focusing on the video "Wonderland Indonesia Chapter 1" with the ID aKtb7Y3qOck, released on Aug 17, 2021. This video has garnered a substantial viewership of 56,636,162 and attracted 246,703 comments. This analysis aims to glean insights into the audience engagement patterns, sentiment trends, and network dynamics surrounding this prominent piece of content. By delving into these aspects, valuable insights are garnered to inform strategic decision-making and content optimization strategies within the creative industry landscape.

![Figure 3: Post-Per-Day and Top 10 Poster (Communalytic)](image)

Figure 3 shows the post-per-day statistics and top ten posters. Based on the data derived from post-per-day statistics, it is discernable that the video mentioned above received 299 comments on Aug 17, 2022, and 220 comments on Aug 18, 2022, indicating a consistent level of engagement over the specified timeframe. Conversely, the engagement pattern for the second video displayed fluctuating activity levels, with 3621 comments recorded on Aug 21, 2022, and 3425 comments on Aug 20, 2022. This variance in engagement levels underscores the importance of analyzing temporal dynamics in understanding audience interactions with digital content, highlighting the need for tailored strategies to sustain engagement and optimize audience reach across different platforms and content types.

Based on the statistics of the top ten posters, it is evident that @yordiantobeny8615 has emerged as the most prolific contributor with 15 posts, followed by @mailina7545 and @NataliSigalingging with six posts each. This distribution underscores the significance of individual user activity in shaping the discourse and content dissemination within the analyzed social network. Furthermore, the varying contribution levels among users highlight the diversity of voices and perspectives contributing to the richness of interactions within digital communities. Such insights from user engagement data offer valuable opportunities for understanding network dynamics and tailoring engagement strategies to foster inclusive and vibrant online communities.

The utilization of data of posts per day and top ten poster statistics at the business understanding stage of the CRISP-DM offers invaluable insights for Social Network Analysis (SNA) and topic analysis. By examining post-per-day statistics, this research discerns patterns of temporal engagement, identifying peak activity periods and trends in audience interaction with digital content. Meanwhile, analysis of the top ten posters provides a deeper understanding of user behavior and influence within the network, shedding light on crucial contributors and their impact on content dissemination. Integrating these data-driven insights into the CRISP-DM framework enhances the foundation for subsequent analytical processes; this research develops more informed strategies and hypotheses to guide their SNA and topic analysis endeavors effectively.

2.2.2 Data Understanding

During the data understanding stage, a meticulous process is undertaken to identify frequently used words by creating word clouds and emoji clouds derived from video review data. This systematic approach aids in uncovering prevalent themes, sentiments, and patterns within the dataset, offering valuable insights into audience preferences, reactions, and engagement with the content. This research comprehensively understands the linguistic and emotive landscape surrounding the video content by leveraging these visualizations, laying a solid groundwork for subsequent analysis and interpretation. Ultimately, this methodological step enhances the depth and breadth of insights derived from the data, contributing to more informed decision-making and strategic content optimization efforts.

Utilizing data word clouds and emoji clouds during the data understanding stage within the Cross-Industry Standard Process for Data Mining (CRISP-DM) methodology offers invaluable benefits for Social Network Analysis (SNA) and topic analysis. By visually representing frequently used words and emojis within the dataset, this research quickly identifies prevalent themes, sentiments, and patterns among users, facilitating a more profound comprehension of audience preferences and reactions to the content. This enhanced understanding is a solid foundation for subsequent analytical processes; this research formulates more targeted hypotheses, identifies key influencers within social networks, and uncovers latent topics of interest. Integrating data word clouds and emoji clouds into the CRISP-DM framework enriches the analytical process, leading to more robust insights and informed decision-making in SNA and topic analysis endeavors.
Calculating frequently used words from the dataset reveals significant linguistic patterns and thematic preferences among viewers. Words such as "Indonesia" occur 176 times, "banget" 116 times, and "lagu" 95 times, indicating a strong association with Indonesian cultural identity and music-related content. Moreover, the repetition of terms like "keren" (82 times) and "bagus" (46 times) underscores the audience's positive reception towards the video, suggesting a prevalent sentiment of admiration and approval. These findings, along with the occurrences of other terms, such as "anak" (62 times), "jangi" (57 times), and "nya" (55 times), provide valuable insights into audience preferences and sentiment trends, offering opportunities for content creators and analysts to tailor their strategies to better resonate with the target audience. Ultimately, this analysis, enriched by numerical occurrences, enhances our understanding of viewer engagement and content dynamics within the digital landscape.

Calculating frequently used emojis from the dataset unveils notable trends in audience expression and sentiment. Emojis such as ❤️ (194 occurrences), ❤️❤️ (54 occurrences), and ❤️ (52 occurrences) emerge as prominent symbols reflecting themes of love, national pride, and celebration within the discourse surrounding the content. Additionally, emoticons like ❤️ (52 occurrences) and ❤️ (42 occurrences) signify emotions ranging from joy to sorrow, underscoring the depth of audience engagement and emotional resonance with the video material. These findings, along with the occurrences of other emojis like 🎉 (24 occurrences) and 🎉 (18 occurrences), provide valuable insights into the nuances of audience reactions and perceptions, enriching our understanding of viewer engagement within digital contexts.

2.2.3 Modeling

Global Vectors for Word Representation (GloVe) is employed during the modeling stage for topic analysis. This choice of methodology underscores the significance of leveraging advanced techniques in natural language processing to extract meaningful insights from textual data. By utilizing GloVe, this research effectively captures semantic relationships between words, enabling a comprehensive understanding of the underlying themes and topics present within the dataset. This approach facilitates the identification of critical topics and their respective distributions, laying a solid foundation for subsequent analysis and interpretation. Ultimately, integrating GloVe
into the modeling process enhances the depth and accuracy of topic analysis; this research uncovers valuable insights and contributes meaningfully to scholarly discourse. The equation used by GloVe to learn word embeddings involves optimizing the following objective function:

\[
J = \sum_{i,j=1}^{V} f(P_{ij})(w_i^T \overline{w}_j + b_i + \overline{b}_j - \log(P_{ij}))^2
\]

Where:
- \( J \) is the objective function to minimize.
- \( V \) is the vocabulary size.
- \( P_{ij} \) is the probability of word \( j \) appearing in the context of word \( i \).
- \( w_i \) and \( \overline{w}_j \) are the word vectors for words \( i \) and \( j \) respectively.
- \( b_i \) and \( \overline{b}_j \) are bias terms for words \( i \) and \( j \) respectively.
- \( f \) is a weighting function, typically \( f(x) = \min \left( \left( \frac{x}{x_{\text{max}}} \right)^{\alpha} , 1 \right) \), where \( x_{\text{max}} \) is a maximum weighting value and \( \alpha \) is a weighting exponent.

Subsequently, a social network analysis (SNA) is conducted, employing a systematic approach to examine the interconnectedness and structural patterns within the social network under study. This analytical method provides a robust framework for understanding the relationships, interactions, and influence dynamics among individuals or entities within the network. By applying SNA, this research uncovers valuable insights into network topology, centrality measures, and community structures, thus enriching our understanding of social phenomena and facilitating strategic interventions for community management or network optimization. Ultimately, integrating SNA into the analytical process enhances the depth and breadth of insights derived from the data, offering valuable opportunities for informed decision-making and effective network management strategies. The equation used in Social Network Analysis (SNA) varies depending on the calculated metric or measure. In addition, standard equations used in SNA include:

\[
C_{\text{degree}}(i) = \frac{\text{Number of connections to node } i}{\text{Total number of nodes} - 1}
\]

\[
C_{\text{betweenness}}(i) = \frac{\sigma_{st}(i)}{\sigma_{st}}
\]

\[
C_{\text{closeness}}(i) = \frac{1}{\sum_{j} d(i,j)}
\]

Where:
- \( C \) denotes the centrality measure.
- \( i \) represents the node in the network.
- \( d(i,j) \) denotes the shortest path between nodes \( i \) and \( j \).
- \( \sigma_{st} \) is the total number of shortest paths between nodes \( s \) and \( t \).
- \( \sigma_{st}(i) \) represents the number of shortest paths between nodes \( s \) and \( t \) that node \( i \) lies on.

Based on the equations elucidated previously, it becomes evident that equation two is utilized to compute degree centrality (2), equation third is employed for calculating betweenness centrality (3), and equation fourth is utilized for determining closeness centrality within a social network analysis context (4). This structured approach to centrality measurement underscores the methodical nature of analyzing network structures, wherein different metrics serve distinct purposes in understanding the significance and influence of nodes within the network. This research systematically quantifies various aspects of node centrality by employing these equations, enriching our understanding of network dynamics and facilitating nuanced interpretations of social phenomena within complex networks.

2.2.4 Evaluation

During the evaluation stage, the evaluation of Global Vectors for Word Representation (GloVe) on the dataset utilized in this research underscores its effectiveness in capturing semantic relationships and contextual information embedded within textual data. Through rigorous assessment, GloVe demonstrates its capability to generate high-quality word embeddings that accurately represent the underlying meanings and nuances of words within the dataset. This evaluation reaffirms GloVe's utility as a powerful tool for natural language processing tasks, providing valuable insights into the semantic structure of the dataset and facilitating more nuanced analyses and interpretations. Consequently, the successful evaluation of GloVe further solidifies its standing as a reliable and effective method for word representation in diverse research contexts.

The Social Network Analysis (SNA) evaluation of the dataset under study reveals valuable insights into the social network's structural dynamics and relational patterns. This research comprehensively understands the network's topology and implications for information flow, influence dynamics, and community formation through rigorous analysis and interpretation of centrality measures, community structures, and network metrics. This holistic evaluation enhances our understanding of the dataset's social fabric and provides actionable insights for strategic interventions, community management, and network optimization. Thus, the evaluation of SNA serves...
as a crucial step in unraveling the complexities of social networks and informing evidence-based decision-making processes.

2.2.5 Deployment

The deployment process of topic analysis results based on review data from video content entails the practical implementation of identified themes and insights to optimize content strategies and enhance audience engagement. By leveraging the findings from the topic analysis, content creators and digital marketers tailor their video content to better resonate with audience interests and preferences, thus increasing viewer satisfaction and retention. This strategic utilization of topic analysis outcomes improves the relevance and effectiveness of video content and fosters a deeper understanding of viewer preferences and sentiments. Consequently, the deployment of topic analysis findings serves as a crucial step in optimizing content creation processes, driving audience engagement, and ultimately maximizing the impact of video content on digital platforms.

The deployment process of Social Network Analysis (SNA) results based on review data from video content involves the practical application of network insights to optimize content strategies and enhance audience engagement within digital platforms. By leveraging the structural patterns, centrality measures, and community structures identified through SNA, content creators and digital marketers tailor their strategies to foster collaborative relationships among viewers, influencers, and content creators, thereby maximizing the reach and impact of video content. This strategic utilization of SNA outcomes enhances audience engagement and fosters a deeper understanding of social dynamics and interactions within digital communities. Consequently, deploying SNA findings is critical in leveraging network insights to drive positive change, foster innovation, and strengthen relationships within digital platforms.

3. RESULT AND DISCUSSION

The discussion within the results of this research comprises the implementation outcomes of the GloVe model in topic analysis, along with the findings from Social Network Analysis (SNA). By integrating the GloVe model into the topic analysis process, this research gained insights into semantic relationships and contextual nuances within textual data, enhancing the depth and accuracy of topic identification. Additionally, the results from SNA provided valuable insights into the structural patterns, centrality measures, and community structures within the analyzed social network, shedding light on the dynamics of interactions and influence among network nodes. This comprehensive discussion underscores the multifaceted approach employed in the research, allowing for a holistic understanding of textual content and social network dynamics, thereby enriching the overall insights derived from the study.

3.1 Topic Analysis: Implementation of Global Vectors for Word Representation (GloVe) Model

The topic analysis model employed for the "Wonderland Indonesia Chapter 1" dataset, identified by the ID aKtb7Y3qOck, is The Global Vectors for Word Representation (GloVe). This model, renowned for its effectiveness in capturing semantic relationships among words based on co-occurrence statistics, offers a robust framework for analyzing textual data. By leveraging GloVe, this research uncovers underlying themes, sentiments, and discourse patterns within the dataset, facilitating a comprehensive understanding of the content's semantic landscape. Consequently, utilizing GloVe enhances the efficiency and accuracy of topic analysis, enabling this research to extract meaningful insights from complex textual data sets such as "Wonderland Indonesia Chapter 1".

The Global Vectors for Word Representation (GloVe) model facilitates topic analysis based on video review data, offering a robust framework for extracting and interpreting thematic patterns embedded within textual content. By harnessing GloVe embeddings, this research effectively captures words' semantic nuances and contextual meanings, enabling a comprehensive understanding of the underlying topics discussed in video reviews. This methodological approach enhances the efficiency and accuracy of topic analysis and enables this research to uncover valuable insights into viewer preferences, sentiments, and perceptions regarding video content. Thus, integrating GloVe into topic analysis methodologies significantly enhances analytical capabilities, fostering deeper insights and informed decision-making in video content analysis.

![Figure 7. Topic Modeling based on Sentiment Classification Vader (Communalytic)](image)
Figure 7 depicts a visualization of topics based on sentiment analysis results using Vader with the Ylorrd color scheme across 23,904 posts. The visualization employs six classifications to determine colors, namely 0, 0.2, 0.4, 0.6, 0.8, and 1, where lighter shades correspond to smaller values and darker shades approach 1. This color scheme facilitates the interpretation of sentiment polarity across the analyzed posts, enabling this research to discern varying degrees of positivity or negativity within the topics under investigation. Ultimately, the visualization provides a comprehensive overview of sentiment distributions, aiding in exploring and understanding sentiment dynamics within the dataset.

![Figure 7: Visualization of topics based on sentiment analysis results using Vader with the Ylorrd color scheme across 23,904 posts.](image)

**Figure 8.** The Global Vectors for Word Representation (GloVe) Model (Communalytic)

Figure 8 illustrates a visualization of topics based on the results of toxicity analysis, encompassing metrics such as toxicity, severe toxicity, profanity, identity attack, insult, and threat, utilizing the Ylorrd color scheme across 23,904 posts. The visualization employs six classifications to determine colors, ranging from 0 to 1, where lighter shades correspond to smaller values and darker shades approach 1. This color scheme facilitates the interpretation of toxicity levels across the analyzed posts, allowing for the identification of varying degrees of inappropriate or harmful content within the topics under investigation. Consequently, the visualization offers valuable insights into the prevalence and distribution of toxic behaviors within the dataset, aiding this research in understanding and addressing issues related to online toxicity effectively.

Topic modeling seamlessly integrated with the results of sentiment classification using Vader, which processed 2204 out of 24185 posts and yielded sentiment scores of 1369 positive (92 or 6.72%), 427 negative (31.19%), and 850 neutral (62.09%). This integration offers a comprehensive approach to understanding textual data by simultaneously uncovering underlying themes and sentiment orientations. Combining topic modeling with sentiment analysis allows this research to gain deeper insights into the content’s semantic landscape and emotional resonance, enabling a nuanced understanding of public opinion and discourse dynamics. Consequently, this integrated approach enhances the efficiency and accuracy of textual data analysis, facilitating informed decision-making processes in various domains, from market research to social media monitoring.

Furthermore, the visualization of topic analysis integrated with the results of sentiment classification using TextBlob, which processed 2204 out of 24185 posts and yielded sentiment scores of 1369 positive (40 or 2.92%), 599 negative (43.75%) and 730 neutral (53.32%). This integration offers a holistic approach to comprehending textual data by concurrently exploring thematic content and sentiment polarity. By merging topic visualization with sentiment analysis, this research gains more profound insights into the semantic and emotional dimensions of the data, facilitating a more nuanced understanding of public opinion and discourse dynamics. Consequently, this integrated approach enhances the comprehensiveness and effectiveness of textual data analysis, empowering informed decision-making across various domains, from market research to social media analytics.

The performance of the Global Vectors for Word Representation (GloVe) model in topic analysis on the review data of the video content "Wonderland Indonesia" by Alffy Rev ft. Novia Bachmid (Chapter 1) demonstrates its efficacy in capturing semantic relationships and thematic patterns within the textual content. By leveraging GloVe embeddings, this research discnerns nuanced topics and themes discussed in the video reviews with precision and granularity. This methodological approach enhances the depth and accuracy of topic analysis, enabling a comprehensive understanding of viewer sentiments, preferences, and perceptions regarding the video content.
content. Consequently, utilizing GloVe in topic analysis facilitates informed decision-making processes and content optimization strategies, thereby enriching the analytical capabilities and insights derived from the study.

3.2 Social Network Analysis (SNA)

The Social Network Analysis (SNA) model in this study visually represents network patterns in the social interactions of YouTube users, particularly those engaging with the video content "Wonderland Indonesia" by Alffy Rev ft. Novia Bachmid (Chapter 1). By leveraging SNA, this research analyzes the user network's structural dynamics, connectivity, and influence patterns, elucidating key nodes, communities, and information flow pathways. This methodological approach enhances our understanding of the complex social dynamics inherent in YouTube interactions, providing valuable insights into user engagement, community formation, and content dissemination strategies. Consequently, integrating SNA into the research framework enriches the analytical depth and facilitates a comprehensive examination of social interactions within the YouTube ecosystem surrounding the specified video content.

Based on the Social Network Analysis (SNA) calculation process, 21,108 actor nodes and 317 edges were observed within the network. This substantial number of actor nodes and edges signifies the complexity and interconnectedness of the analyzed social network. Such extensive connectivity among actor nodes suggests a rich web of relationships and interactions, which provide valuable insights into the structure and dynamics of the studied social system. Consequently, this comprehensive analysis facilitated by SNA offers a nuanced understanding of social relationships and their implications, serving as a foundational tool for various disciplines ranging from sociology to marketing.

Figure 9 shows the SNA visualization based on number of nodes. Based on the node size measured by in-degree centrality, three main clusters emerge, each delineated by the number of nodes: 38, 24, and 21, respectively. This segmentation based on node size underscores the presence of distinct groupings within the network, each characterized by varying degrees of connectivity and influence. Identifying these clusters provides valuable insights into the structural organization of the network, highlighting key nodes and their significance in facilitating information flow and interaction dynamics. Consequently, this analysis aids in understanding the intricate interplay of nodes within the network, offering a nuanced perspective on its underlying structure and functioning.

Figure 10. Indegree Centrality (Atlas Nomic)
Figure 10 shows the indegree centrality using atlas nomic. The purpose of visualizing Social Network Analysis (SNA) based on sentiment analysis of video content is to elucidate the intricate relationship between social interactions and emotional responses within online communities. By integrating sentiment analysis with SNA, this research aims to uncover how sentiments expressed by individuals influence the structure and dynamics of social networks surrounding video content. This amalgamation of techniques enables a deeper understanding of the underlying sentiments driving user engagement and interaction patterns, thereby facilitating targeted interventions to enhance content dissemination strategies and community engagement initiatives. Consequently, visualizing SNA through sentiment analysis offers a holistic approach to deciphering the complex interplay between social networks and emotional dynamics within video content.

4. CONCLUSION

This study examines the reception and impact of the "Wonderland Indonesia" video by Alfify Rev ft. Novia Bachmid (Chapter 1) within the YouTube community, employing a comprehensive methodology based on the CRoss-Industry Standard Process for Data Mining (CRISP-DM), topic analysis, and Social Network Analysis (SNA). The study identifies the content's main themes and narrative elements through topic analysis, shedding light on its effectiveness in storytelling. Furthermore, sentiment analysis using Vader and TextBlob reveals varying levels of sentiment polarity across posts, with 1369 (92%) of the posts analyzed using Vader exhibiting positive sentiment, 427 (31.19%) neutral sentiment, and 850 (62.09%) negative sentiment. Similarly, sentiment analysis using TextBlob on the same subset of posts shows 1369 (40) posts with positive sentiment, 599 (43.75%) neutral sentiment, and 730 (53.32%) expressing negative sentiment. Notably, metrics such as toxicity and severe toxicity exhibit varying degrees of prominence within the content, with the highest values being 0.90780 and 0.95021, respectively. These findings inform targeted interventions and content moderation strategies to foster healthier online discourse. The SNA uncovers intricate social dynamics and interaction patterns among viewers, highlighting the video's ability to facilitate engagement and community interaction. This research underscores the importance of creative storytelling and community engagement strategies in digital content creation, with implications for audience participation and community development in the digital realm. Future research avenues may explore the longitudinal effects of such strategies on audience retention and community engagement.

ACKNOWLEDGMENT

Acknowledgment to the Tourism Department, Faculty of Business Administration and Communication, LPPM, Atma Jaya Catholic University of Indonesia, is extended with utmost gratitude for invaluable support and resources throughout this research endeavor.

REFERENCES


Copyright © 2024 Yerik Afrianto Singgalen, Page 569
This Journal is licensed under a Creative Commons Attribution 4.0 International License