International Journal on Social Science, Economics and Art, 13 (4) (2024) 255-267

Published by: Institute of Computer Science (IOCS)

International Journal on Social Science, Economics and Art

Journal homepage: www. ijosea.isha.or.id

Social network and sentiment analysis of product reviews (case of smartwatch product content)

Yerik Afrianto Singgalen[1]

[1,] Faculty of Business Administration and Communication, Atma Jaya Catholic University of Indonesia, Indonesia

Article Info

Article history:

Received: Jan 17, 2024
Revised: Feb 18, 2024
Accepted: Feb 28, 2024

Keywords:

Sentiment;
Social Network;
Product;
Smartwatch;
Reviews.

ABSTRACT

This study addresses the need to understand the dynamics of sentiment and social network analysis (SNA) in the context of smartwatch product reviews. Leveraging the CRoss-Industry Standard Process for Data Mining (CRISP-DM) methodology, the research aims to analyze sentiments and social networks to glean insights into consumer behavior and interaction patterns. The CRISP-DM framework guides the research through structured phases of business understanding, data understanding, data preparation, modeling, evaluation, and deployment. Through sentiment analysis using Support Vector Machine (SVM) with Synthetic Minority Over-sampling Technique (SMOTE) and SNA, the study examines accuracy (91.41% +/- 1.66%), precision (100.00% +/- 0.00%), recall (82.80% +/- 3.36%), f-measure (90.56% +/- 2.01%), Area Under the Curve (AUC), as well as network metrics such as diameter (4), density (0.001036), reciprocity (0.000000), centralization (0.004920), and modularity (0.994200). Findings reveal a robust performance of the SVM algorithm coupled with SMOTE, showcasing high accuracy and effective discrimination between sentiments. Additionally, SNA uncovers valuable insights into network structures, communication patterns, and sentiment propagation dynamics within the online community. These findings contribute to a deeper understanding of consumer sentiments and interactions, guiding strategic marketing, product development, and reputation management decisions.

Corresponding Author:

Yerik Afrianto Singgalen
Faculty of Business Administration and Communication,
Atma Jaya Catholic University of Indonesia
Jl. Jend. Sudirman No.51, RT.5/RW.4, Karet Semanggi, Kecamatan Setiabudi, Kota Jakarta Selatan, Daerah Khusus Ibukota Jakarta 12930.
Email: yerik.afrianto@atmajaya.ac.id

1. INTRODUCTION

Influencer product reviews have emerged as a prominent tool in shaping consumer purchasing intentions. These reviews, often disseminated through various social media platforms, wield significant influence due to their ability to reach vast audiences and convey authentic experiences with products (Fang et al., 2022). By leveraging their expertise and credibility within specific niches, influencers offer nuanced evaluations that resonate with consumers seeking informed purchasing decisions (Botchway et al., 2020). Consequently, these reviews are pivotal in shaping consumer perceptions and preferences, ultimately driving purchasing behavior (Rodrigues et al., 2020). As such, collaborating with influencers and brands in delivering product reviews is a strategic approach to engage and persuade consumers in today's digital landscape.

Digital content featuring product reviews has proven to be a catalyst in stimulating consumer purchasing interest and shaping sentiments toward brands. Through comprehensive evaluations and firsthand experiences shared by influencers or content creators, such content is valuable for consumers seeking guidance in their purchasing decisions (Bilianos, 2022; Güneş, 2020). The informative nature of these reviews educates consumers about product features and functionalities and helps build trust and credibility around the brand (Suri et al., 2023; van Atteveldt et al., 2021). As consumers engage with these reviews, they develop a deeper understanding of the brand's offerings, fostering positive sentiments and potentially leading to brand loyalty (Celuch, 2021). Consequently, the strategic integration of product reviews into digital content plays a pivotal role in influencing consumer behavior and establishing enduring relationships between consumers and brands.

The urgency of this research lies in its potential to address critical gaps in understanding and inform strategic decision-making in the digital marketing landscape. With the exponential growth of social media platforms and the increasing influence of online consumer sentiments on brand perception and purchasing behavior, there is a pressing need for comprehensive analyses to unravel the intricate dynamics of social network interactions and sentiment classification(Vanderkooi et al., 2023). By elucidating these phenomena, this research contributes to academic knowledge and provides actionable insights for marketers seeking to navigate the complexities of digital engagement and capitalize on emerging trends (Lappeman et al., 2023). Thus, the timeliness of this research is paramount in guiding effective strategies and fostering sustainable competitive advantages in an ever-evolving digital ecosystem.

This study analyzes the social network and consumer sentiment regarding smartwatch products within video content identified by the ID 2sqIRlHR4Mg. By employing advanced data mining techniques, researchers intend to map out the interconnectedness of users engaging with this video and decipher the prevailing sentiments expressed towards smartwatch products (Kumar et al., 2023; Rahimi et al., 2022; Ray & Chakrabarti, 2022; Sarmast et al., 2023). Through comprehensive sentiment analysis algorithms, the study seeks to uncover patterns and trends in consumer opinions, providing valuable insights into consumer perceptions and preferences (Areni, 2022; Hossain et al., 2023; Oh & Yi, 2022; Rajeswari et al., 2020; Srivastava et al., 2022; Yerpude & Rautela, 2022). Ultimately, this research contributes to a deeper understanding of consumer behavior within digital media consumption. It offers actionable insights for marketers and industry stakeholders aiming to optimize their strategies in the smartwatch market.

Theoretical and practical implications of this research hold significant importance in advancing academic understanding and real-world applications within the studied domain. By elucidating the underlying mechanisms and dynamics of social network interactions and consumer sentiment analysis, this study contributes to the theoretical frameworks within digital marketing and consumer behavior research (Hoskins et al., 2021). Furthermore, the practical implications of this research extend to industry practitioners and marketers, offering actionable insights into optimizing digital marketing strategies and enhancing brand engagement in the era of social media influence (La et al., 2022; Lappeman et al., 2022; Mehmet & D'Alessandro, 2022). Consequently, this research bridges the gap between theory and practice, fostering knowledge transfer and facilitating informed decision-making processes in both academic and commercial spheres.

This research employs the CRISP-DM methodology to analyze social networks and classify sentiment, enhancing the study's systematic and efficient examination of data. By utilizing this well-established framework, researchers can navigate through the various stages of data mining, from understanding business objectives to evaluating model performance, thereby ensuring robustness and reliability in the analysis process (Singgalen, 2023c, 2023a, 2023e, 2023f). The adoption of CRISP-DM facilitates a structured approach to data analysis and enables researchers to effectively communicate findings and insights to academic and industry audiences (Singgalen, 2023d, 2023g, 2023b). Consequently, the utilization of CRISP-DM methodology underscores the rigor and credibility of the research, contributing to its overall validity and applicability in addressing contemporary challenges within the field of digital media analysis.

Exploring similar research and acknowledging limitations are integral to scholarly discourse within any field. While existing studies may have investigated related topics or utilized comparable methodologies, each research endeavor possesses unique contextual nuances and objectives. Therefore, while drawing insights from similar research can provide valuable reference points and theoretical foundations, it is essential to recognize the distinctive contributions and advancements offered by the current study (Pfeuffer et al., 2021; Wei Lun Lee et al., 2023). Furthermore, addressing the limitations inherent in the research process, such as sample size constraints or data collection challenges, ensures transparency and integrity in scholarly inquiry (Dash et al., 2021; Zhu et al., 2023). Scholars can foster a nuanced understanding of the research landscape by critically evaluating similar research and limitations, thereby advancing knowledge and facilitating continuous improvement in academic endeavors.

2. RESEARCH METHOD

The CRISP-DM (Cross-Industry Standard Process for Data Mining) methodology is a systematic and efficient framework for conducting data mining projects. It encompasses six significant phases, including business understanding, data understanding, data preparation, modeling, evaluation, and deployment, each contributing distinctively to the overall success of the analytical process. With its structured approach, CRISP-DM enables researchers and practitioners to navigate the complexities of data analysis, ensuring methodological rigor and comprehensive exploration of insights. Its widespread adoption across various industries underscores its versatility and effectiveness in addressing diverse analytical challenges. Thus, the CRISP-DM methodology is a cornerstone in modern data mining practices, providing a robust foundation for informed decision-making and actionable insights.

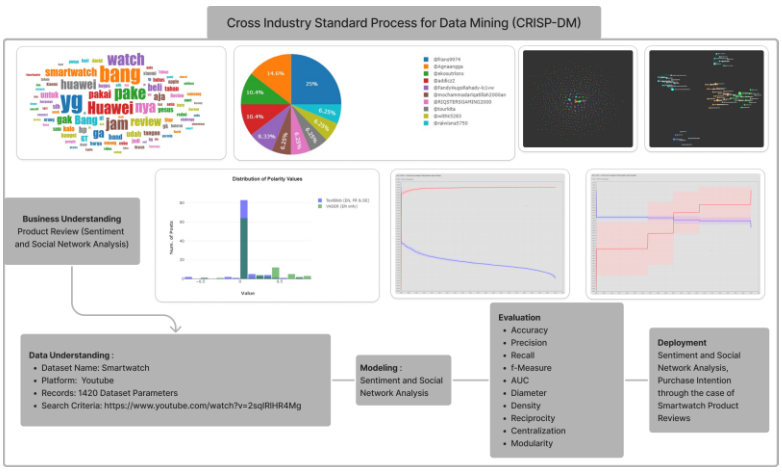

Figure 1. Cross Industry Standard Process for Data Mining (CRISP-DM)

The CRISP-DM (Cross-Industry Standard Process for Data Mining) methodology exhibits several advantages contributing to its prominence in data mining projects. Its systematic structure, comprising distinct phases such as business understanding, data preparation, modeling, evaluation, and deployment, provides a clear roadmap for researchers and practitioners, ensuring methodological rigor and efficiency throughout the analytical process. Moreover, CRISP-DM's iterative nature allows for flexibility and adaptation to evolving project requirements, facilitating continuous improvement and refinement of analytical models. Consequently, CRISP-DM stands out as a robust and adaptable framework, empowering analysts to navigate complex data landscapes and derive actionable insights effectively, thereby enhancing decision-making processes and driving organizational success.

2.1 Business Understanding

During the business understanding phase, the contextual information regarding smartwatch product reviews within the video identified by the ID 2sqIRlHR4Mg is analyzed. This initial stage entails a comprehensive understanding of the project's objectives, requirements, and constraints, explicitly focusing on the pertinent aspects of smartwatch products and consumer sentiment as portrayed in the specified video content. By delving into the nuances of the smartwatch market and consumer preferences highlighted within the video, researchers can lay a solid foundation for subsequent data preparation and analysis stages, thereby ensuring the relevance and effectiveness of the analytical process. Consequently, the meticulous examination of the business context surrounding the smartwatch product reviews facilitates informed decision-making and enhances the overall success of the research endeavor.

Figure 2. Post Per Day (Communalityc)

Based on the data indicating the number of posts per day, it is evident that there was a substantial increase in activity on October 3, 2023, with 610 posts compared to the relatively lower activity of 85 posts on October 4, 2023. This disparity suggests a significant engagement or content generation fluctuation within the specified time frame, possibly influenced by external factors such as trending topics, events, or promotional campaigns. Such fluctuations in posting activity underscore the dynamic nature of online communities and the potential impact of various contextual factors on user participation and content creation behaviors. Consequently, analyzing patterns in post frequency can offer valuable insights into the temporal dynamics of online discourse and inform strategies for community management or content moderation.

2.2 Data Understanding

During the data understanding phase, identifying the top ten posters is crucial in discerning the users who actively engage with and potentially influence the viewership of product reviews within the video content. By analyzing the posting behavior and frequency of these top contributors, researchers can gain insights into the key influencers within the community and their impact on shaping discussions or opinions regarding the featured products. This strategic approach enables a targeted examination of user interactions, facilitating a more nuanced understanding of the dynamics surrounding product reviews and audience engagement within the specified video context. Consequently, identifying top posters enhances the comprehensiveness of data understanding and lays the groundwork for subsequent analyses to decipher user behavior and sentiments effectively.

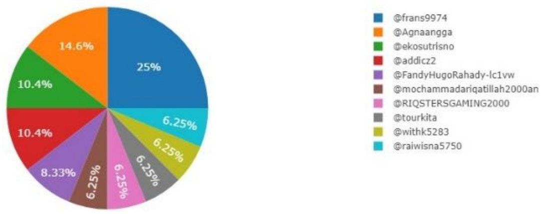

Figure 2. Top Ten Poster (Communalityc)

The data on the top ten posters reveals distinct patterns in user engagement within the context of the product review video. Notably, @frans9974 emerges as the most prolific contributor with 12 posts, suggesting a substantial and consistent level of participation. Contributions such as @Agnaangga, @ekosutrisno, and @addicz2 demonstrate notable engagement with 7, 5, and 5 posts, respectively. This distribution of posting activity highlights the existence of critical influencers and active participants within the community, indicating potential nodes of influence that can significantly impact the discourse surrounding product reviews. Identifying these top posters enriches the understanding of user behavior and sets the stage for targeted analyses to decipher their impact on shaping sentiments and opinions within the online community. Consequently, this detailed examination of top posters provides a valuable lens through which researchers can navigate the intricate dynamics of user participation and influence in the digital discourse.

2.3 Modeling

During the modeling phase, a dual analysis involving sentiment analysis and social network analysis is undertaken. This critical stage encompasses the application of sophisticated algorithms to evaluate the sentiment expressed within the user-generated content, offering a nuanced understanding of the collective attitudes toward the reviewed products. Simultaneously, the social network analysis unveils the intricate connections and interactions among users, shedding light on the structural patterns of the online community. Combining sentiment and social network analysis, this multi-faceted approach provides a comprehensive framework for deciphering the interplay between user sentiments and the underlying social dynamics. By concurrently examining the emotional responses and network structures, researchers can derive more profound insights into the factors influencing online discourse and the potential impact of influential users on shaping sentiments within the digital community. In conclusion, this dual analysis at the modeling phase enhances the robustness of the research, fostering a holistic understanding of the complex dynamics present in the online environment.

Figure 3. Frequently Used Words (Communalityc)

Identifying frequently used words reveals significant insights into the thematic focus and prevalent topics within the discourse surrounding the reviewed product. Among the top frequently used words, "bang" appears 145 times, followed by "pake" with 113 occurrences, "Huawei" with 108 mentions, "watch" with 106 instances, and "jam" with 102 appearances. The repeated occurrence of these terms underscores their significance in shaping the narrative and highlights critical areas of interest among users. Additionally, the presence of terms like "review," which occurs 77 times, suggests an inclination towards evaluating and discussing the features and performance of the Huawei smartwatch. This detailed analysis of frequently used words provides valuable insights into the dominant themes and topics driving the conversation, informing subsequent analyses and facilitating a deeper understanding of user sentiments and preferences regarding the product. Consequently, leveraging these insights can guide strategic decision-making and enhance engagement strategies within the digital discourse surrounding the Huawei smartwatch.

2.4 Evaluation

During the evaluation phase, a comprehensive analysis of the effectiveness of digital marketing is conducted by synthesizing influencer content with data derived from Social Network Analysis (SNA) and sentiment analysis. This critical stage involves assessing the impact and reach of influencer-led campaigns, considering the network dynamics and sentiments expressed by the audience. By triangulating influencer-generated content with SNA metrics and sentiment insights, researchers can gauge the resonance of digital marketing efforts within the target audience. This multifaceted approach not only offers a quantitative understanding of reach and engagement but also provides qualitative insights into the sentiments and perceptions of the audience. Consequently, integrating SNA and sentiment analysis into the evaluation process enhances the comprehensiveness of assessing digital marketing effectiveness, providing valuable data-driven insights for refining strategies and maximizing the impact of influencer-led campaigns.

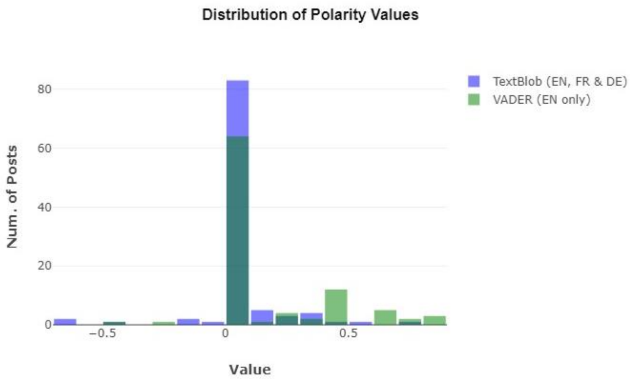

Figure 4. Pilot Testing of Sentiment Analysis using Vader (Communalityc)

Based on the pilot testing of sentiment analysis using the VADER (Valence Aware Dictionary and Sentiment Reasoner) tool on a dataset comprising 104 observations, it is discerned that most of the data, 67.37%, exhibits a neutral sentiment. In contrast, 30.53% is attributed to positive sentiment. Conversely, only 2.11% of the data depicts negative sentiment. This distribution underscores the utility of VADER in effectively categorizing sentiment within the dataset, with a predominant focus on neutral expressions followed by positive sentiments. The minimal negative sentiment suggests a generally favorable sentiment toward the subject matter under analysis. Consequently, these findings highlight the efficacy of the VADER sentiment analysis tool in discerning nuanced sentiment patterns within textual data, thereby offering valuable insights for subsequent analyses and decision-making processes.

Based on the results of identifying patterns in Social Network Analysis (SNA), several key metrics provide insights into the structural characteristics of the analyzed network. With a diameter 4, the network demonstrates a relatively short maximum path length between nodes, indicative of efficient communication and connectivity. Furthermore, the low density of 0.001036 suggests sparse connections among nodes, highlighting potential opportunities for network expansion or enhancement of connections. The absence of reciprocity at 0.000000 indicates that connections are predominantly unidirectional, emphasizing asymmetrical relationships within the network. Additionally, the centralization value 0.004920 suggests a decentralized network structure, with power and influence distributed among multiple nodes rather than concentrated in a few. Finally, the high modularity score of 0.994200 signifies the presence of distinct subgroups or communities within the network, indicating specialized clusters of interconnected nodes. Overall, these SNA metrics offer valuable insights into the topology and organization of the network, guiding strategic decisions and interventions aimed at optimizing network performance and functionality.

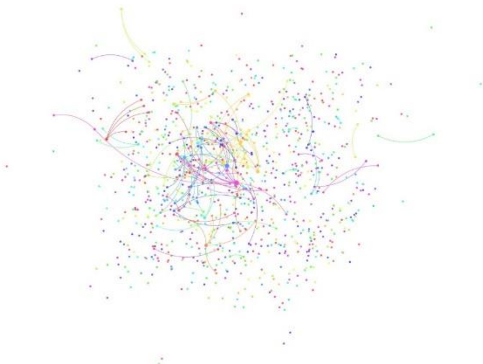

Figure 5. Social Network Analysis of Smartwatch Product Reviews from Video 2sqIRlHR4Mg (Netlytic)

The analysis provided offers a comprehensive understanding of the structural characteristics of the network under study through the lens of Social Network Analysis (SNA). The key metrics, including diameter, density, reciprocity, centralization, and modularity, provide valuable insights into various aspects of the network's organization and functionality. For instance, the relatively short diameter of 4 indicates efficient communication pathways between nodes, facilitating quick dissemination of information. Moreover, the low density suggests the network has sparse connections, presenting opportunities for expanding and strengthening connections. The absence of reciprocity underscores the prevalence of unidirectional relationships, while the decentralized structure, indicated by the centralization value, suggests a balanced distribution of power and influence. Finally, the high modularity score implies the presence of distinct communities within the network, highlighting specialized clusters of interconnected nodes. Overall, these insights guide strategic decision-making processes to enhance network performance and functionality, underscoring the significance of SNA in understanding and optimizing complex network systems.

2.5 Deployment

Purchasing intent can be systematically evaluated during deployment by leveraging sentiment and social network analysis data. This pivotal stage involves translating the insights from these analyses into actionable strategies that can influence and optimize consumer interest. The integration of sentiment analysis enables a nuanced understanding of consumer attitudes toward a product or brand, providing valuable information that can inform targeted marketing efforts. Simultaneously, social network analysis offers insights into the influence dynamics within online communities, identifying key nodes that can impact purchasing decisions. By combining these analytical approaches, marketers can strategically deploy interventions that align with identified sentiments and leverage influential nodes within the network, ultimately guiding and enhancing purchasing intent. Consequently, the deployment phase becomes critical for translating analytical findings into practical applications, ensuring a cohesive and effective strategy for influencing consumer behavior.

3. RESULTS AND DISCUSSIONS

3.1. Sentiment Analysis of The Content Reviews

The urgency of conducting sentiment analysis on smartwatch product reviews is underscored by the dynamic nature of the digital landscape and its profound impact on consumer behavior. As the smartwatch market evolves rapidly, consumer sentiments are pivotal in shaping market trends and influencing purchasing decisions. Analyzing sentiments expressed in product reviews provides a real-time reflection of consumer perceptions and unveils crucial insights into the factors that drive product satisfaction or dissatisfaction. Moreover, sentiment analysis aids in deciphering the overall sentiment polarity, helping companies identify areas of improvement and capitalize on strengths. In a fast-paced market, where consumer preferences can shift swiftly, timely sentiment analysis becomes imperative for businesses to adapt their strategies effectively, stay competitive, and enhance their products to meet evolving consumer expectations. In conclusion, the urgency of sentiment analysis in smartwatch product reviews lies in its capacity to offer actionable insights, enabling companies to stay agile and responsive in a rapidly changing market environment.

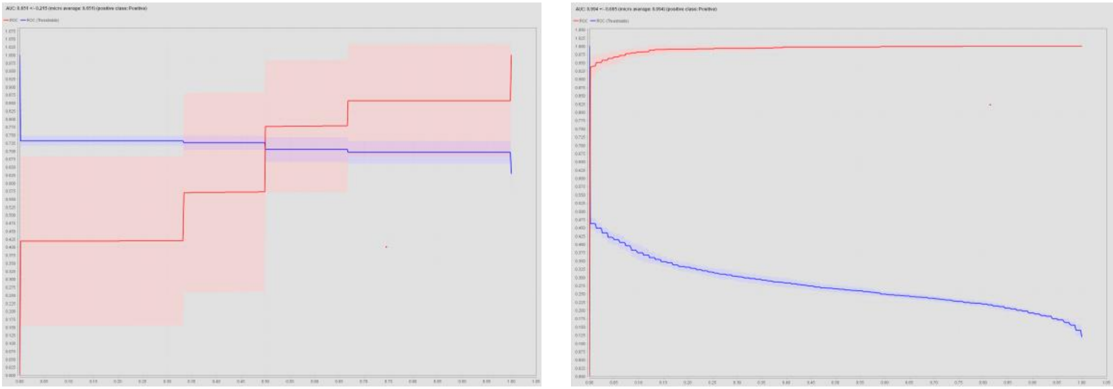

Figure 6. Area Under Curve of SVM with and without SMOTE (Rapidminer)

Based on analyzing the performance of the Support Vector Machine (SVM) algorithm using the Synthetic Minority Over-sampling Technique (SMOTE), insights into the algorithm's efficacy emerge. The PerformanceVector indicates an accuracy of 91.41% +/- 1.66%, showcasing the model's robust performance in correctly classifying instances. The ConfusionMatrix further illustrates the model's proficiency, with a high precision of 100.00% +/- 0.00% and recall of 82.80% +/- 3.36% for the positive class. This precision-recall balance is reflected in the impressive F-measure of 90.56% +/- 2.01%. Additionally, the Area Under the Curve (AUC) values consistently exceeding 0.99 affirm the algorithm's capability to discriminate between classes effectively. The micro average metrics provide a comprehensive overview, emphasizing the SVM's consistency across various performance indicators. In conclusion, the SVM algorithm and SMOTE exhibit commendable performance metrics, affirming its suitability for the given task and suggesting its potential applicability in similar contexts.

Analyzing the Support Vector Machine (SVM) algorithm's performance without employing Synthetic Minority Over-sampling Technique (SMOTE) reveals noteworthy metrics indicative of its proficiency in classification tasks. The PerformanceVector showcases an impressive accuracy of 97.20% +/- 0.61%, underlining the algorithm's ability to classify instances correctly. The ConfusionMatrix, while exhibiting 100% recall for the positive class, demonstrates a precision of 97.20% +/- 0.61%, emphasizing the model's capability to minimize false positives. The Area Under the Curve (AUC) values, albeit lower than those with SMOTE, suggest a satisfactory discriminatory ability of the algorithm. The consistently high f-measure of 98.58% +/- 0.31% further affirms the algorithm's precision-recall balance. Despite the lower AUC values compared to the SMOTE-enhanced model, these results highlight the SVM algorithm's robustness in handling imbalanced datasets without oversampling techniques. In conclusion, even without SMOTE, the SVM algorithm showcases commendable performance metrics, indicating its effectiveness and potential suitability for tasks involving imbalanced datasets.

Analyzing the Support Vector Machine (SVM) algorithm, both with and without the Synthetic Minority Over-sampling Technique (SMOTE), provides valuable insights into its performance in classification tasks. When coupled with SMOTE, the SVM algorithm exhibits commendable metrics, as evidenced by an accuracy of 91.41% +/- 1.66%, a precision of 100.00% +/- 0.00%, and a recall of 82.80% +/- 3.36% for the positive class. The F-measure of 90.56% +/- 2.01% further emphasizes the algorithm's precision-recall balance. Additionally, consistently high Area Under the Curve (AUC) values exceeding 0.99 affirm the SVM's effective class discrimination. The micro average metrics highlight the algorithm's consistency across various performance indicators. On the other hand, even without SMOTE, the SVM algorithm demonstrates noteworthy performance, boasting an accuracy of 97.20% +/- 0.61%, a precision of 97.20% +/- 0.61%, and a recall of 100.00% for the positive class. The F-measure remains impressively high at 98.58% +/- 0.31%. While AUC values are slightly lower than the SMOTE-enhanced model, the results underscore the algorithm's robustness in handling imbalanced datasets without oversampling techniques. In conclusion, the SVM algorithm exhibits effectiveness in classification tasks, with or without SMOTE, suggesting its versatility and potential suitability for various applications involving imbalanced datasets.

3.2. Social Network Analysis: Smartwatch Product Reviews

The urgency of conducting Social Network Analysis (SNA) on Smartwatch Product Reviews is paramount in the contemporary landscape of digital consumerism. In an era where online platforms serve as prominent arenas for product discourse, understanding the intricate social dynamics within these networks is essential. The main topic revolves around deciphering the relationships, influence patterns, and sentiment diffusion among users engaging with smartwatch reviews. This urgency is underscored by the vast volume of user-generated content, making it imperative to unveil the underlying structures and communities shaping the narrative around smartwatches. The significance of this analysis lies in its capacity to reveal not only individual sentiments but also the collective influence of network structures on consumer perceptions. By employing SNA, researchers can unravel critical influencers, identify communication hubs, and discern emergent trends within the interconnected web of smartwatch discussions. In conclusion, the urgency of SNA in Smartwatch Product Reviews is integral to gaining nuanced insights into the evolving landscape of consumer opinions, facilitating informed decision-making for businesses, and enhancing our understanding of the intricate relationships within digital communities.

Based on the results of the Social Network Analysis (SNA) visualization, specifically focusing on "who mentions whom," a comprehensive overview emerges with 1094 posters forming 2192 ties within the threaded discussion dataset. The main topic revolves around the statistical metrics derived from this analysis. The dataset type being a threaded discussion indicates a specific context of interaction. The subsequent metrics, including a Diameter of 0, Density of 0.000915, and Centralization of 0.000000, suggest a network characterized by minimal path lengths, sparse connections, and a decentralized structure. The NaN (Not a Number) value for Reciprocity indicates a lack of bidirectional connections, emphasizing the asymmetrical nature of relationships within the network. The exceptionally high Modularity score of 0.999100 points to the presence of distinct and well-defined subgroups or communities within the network. This systematic exploration of SNA metrics unveils the network's intricate structures and characteristics, providing a nuanced understanding of interaction patterns. In conclusion, these findings offer valuable insights into the complex dynamics of information flow and community formation within threaded discussions, guiding further inquiries into the nature of online interactions.

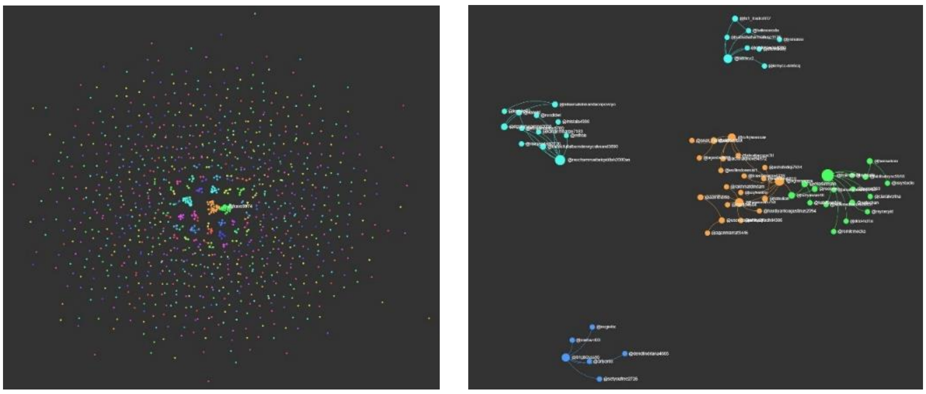

Figure 6. Who Replies to Whom Network (Netlytic)

Based on the results of the Social Network Analysis (SNA) visualization focusing on "who replies to whom," the examination reveals a network comprising 10945 posters connected by 1305 ties. This main topic highlights the statistical metrics derived from the analysis. The subsequent exploration of metrics, including a Diameter of 4 and a Density of 0.001036, indicates a network characterized by relatively short maximum path lengths between nodes and sparse connections among posters. The absence of Reciprocity, represented by a value of 0.000000, suggests predominantly unidirectional interactions within the network. The Centralization value 0.004920 also signifies a decentralized structure, with power and influence distributed across multiple nodes rather than concentrated in a few. The high Modularity score of 0.994200 points to distinct subgroups or communities within the network. This systematic examination of SNA metrics provides valuable insights into the analyzed network's structural characteristics and interaction patterns. In conclusion, these findings contribute to a deeper understanding of the dynamics of communication and engagement among posters, facilitating further research into the complexities of online interactions.

4. CONCLUSION

In conclusion, the integration of Social Network Analysis (SNA) of "who replies to whom" and Sentiment Analysis using a Support Vector Machine (SVM) with Synthetic Minority Over-sampling Technique (SMOTE) presents a comprehensive framework for understanding online interactions and sentiment dynamics. The analysis of SVM performance with SMOTE reveals commendable accuracy of 91.41% +/- 1.66%, high precision of 100.00% +/- 0.00%, recall of 82.80% +/- 3.36%, and F-measure of 90.56% +/- 2.01%, as well as consistent Area Under the Curve (AUC) values exceeding 0.99, indicating effective discrimination between classes. Moreover, the SNA findings indicate a network structure with a diameter of 4, density of 0.001036, reciprocity of 0.000000, centralization of 0.004920, and modularity of 0.994200, along with 10945 posters connected by 1305 ties, offering valuable insights into communication patterns and network characteristics. These insights offer a valuable understanding of the analyzed online community's sentiment dynamics and communication patterns. Consequently, this integrated approach can guide strategic decisions in marketing, reputation management, and product development, suggesting its potential applicability in similar contexts and underscoring its effectiveness in optimizing network performance and functionality while handling imbalanced datasets effectively.

REFERENCES

Areni, C. S. (2022). Automated text analyses of YouTube comments as field experiments for assessing consumer sentiment towards products and brands. *Journal of Product & Brand Management*, *31*(5), 702–717. https://doi.org/10.1108/JPBM-01-2021-3341

Bilianos, D. (2022). Experiments in Text Classification: Analyzing the Sentiment of Electronic Product Reviews in Greek. *Journal of Quantitative Linguistics*, *29*(3), 374–386. https://doi.org/10.1080/09296174.2021.1885872

Botchway, R. K., Jibril, A. B., Oplatková, Z. K., & Chovancová, M. (2020). Deductions from a Sub-Saharan African Bank's Tweets: A sentiment analysis approach. *Cogent Economics and Finance*, *8*(1). https://doi.org/10.1080/23322039.2020.1776006

Celuch, K. (2021). Customers' experience of purchasing event tickets: mining online reviews based on topic modeling and sentiment analysis. *International Journal of Event and Festival Management*, *12*(1), 36–50. https://doi.org/10.1108/IJEFM-06-2020-0034

Dash, A., Zhang, D., & Zhou, L. (2021). Personalized Ranking of Online Reviews Based on Consumer Preferences in Product Features. *International Journal of Electronic Commerce*, *25*(1), 29–50. https://doi.org/10.1080/10864415.2021.1846852

Fang, H., Jiang, G., & Li, D. (2022). Sentiment analysis based on Chinese BERT and fused deep neural networks for sentence-level Chinese e-commerce product reviews. *Systems Science and Control Engineering*, *10*(1), 802–810. https://doi.org/10.1080/21642583.2022.2123060

Güneş, S. (2020). Extracting Online Product Review Patterns and Causes: A New Aspect/Cause Based Heuristic for Designers. *Design Journal*, *23*(3), 371. https://doi.org/10.1080/14606925.2020.1746611

Hoskins, J., Verhaal, J. C., & Griffin, A. (2021). How within-country consumer product (or brand) localness and supporting marketing tactics influence sales performance. *European Journal of Marketing*, *55*(2), 565–592. https://doi.org/10.1108/EJM-11-2018-0787

Hossain, M. S., Begum, H., Rouf, M. A., & Sabuj, M. M. I. (2023). Investigation and prediction of users' sentiment toward food delivery apps applying machine learning approaches. *Journal of Contemporary Marketing Science*, *6*(2), 109–127. https://doi.org/10.1108/JCMARS-12-2022-0030

Kumar, A., Shrivastav, S. K., & Bhattacharyya, S. (2023). Measuring strategic fit using big data analytics in the automotive supply chain: a data source triangulation-based research. *International Journal of Productivity and Performance Management*, *72*(10), 2977–2999. https://doi.org/10.1108/IJPPM-11-2021-0672

La, L., Xu, F., Ren, Q., Zhen, F., & Lobsang, T. (2022). Understanding Consumers' Sentiment Expressions in Online Reviews: A Hybrid Approach. *Journal of International Consumer Marketing*, *34*(4), 465–479. https://doi.org/10.1080/08961530.2021.1982807

Lappeman, J., Goder, A., Naicker, K., Faruki, H., & Gordon, P. (2023). Using Sentiment Analysis to Understand Public Policy Nicknames: Obamacare and the Affordable Care Act. *Journal of Nonprofit and Public Sector Marketing*, *00*(00), 1–17. https://doi.org/10.1080/10495142.2023.2178588

Lappeman, J., Meyer, S., & Miguel, D. (2022). Exploring Online Sentiment (OS) As a Measure of Customer Experience (CX) for Telecommunication Services. *Services Marketing Quarterly*, *43*(3), 257–276. https://doi.org/10.1080/15332969.2021.1992558

Mehmet, M. I., & D'Alessandro, S. (2022). More than words can say: a multimodal approach to understanding meaning and sentiment in social media. *Journal of Marketing Management*, *38*(13–14), 1461–1493. https://doi.org/10.1080/0267257X.2022.2105933

Oh, Y. K., & Yi, J. (2022). Asymmetric effect of feature level sentiment on product rating: an application of bigram natural language processing (NLP) analysis. *Internet Research*, *32*(3), 1023–1040. https://doi.org/10.1108/INTR-11-2020-0649

Pfeuffer, A., Lu, X., Zhang, Y., & Huh, J. (2021). The Effect of Sponsorship Disclosure in YouTube Product Reviews. *Journal of Current Issues and Research in Advertising*, *42*(4), 391–410. https://doi.org/10.1080/10641734.2020.1859023

Rahimi, R., Thelwall, M., Okumus, F., & Bilgihan, A. (2022). Know your guests' preferences before they arrive at your hotel: evidence from TripAdvisor. *Consumer Behavior in Tourism and Hospitality*, *17*(1), 89–106. https://doi.org/10.1108/CBTH-06-2021-0148

Rajeswari, B., Madhavan, S., Venkatesakumar, R., & Riasudeen, S. (2020). Sentiment analysis of consumer reviews – a comparison of organic and regular food products usage. *Rajagiri Management Journal*, *14*(2), 155–167. https://doi.org/10.1108/RAMJ-05-2020-0022

Ray, P., & Chakrabarti, A. (2022). A Mixed approach of Deep Learning method and Rule-Based method to improve Aspect Level Sentiment Analysis. *Applied Computing and Informatics*, *18*(1/2), 163–178. https://doi.org/10.1016/j.aci.2019.02.002

Rodrigues, A. P., Chiplunkar, N. N., & Fernandes, R. (2020). Aspect-based classification of product reviews using Hadoop framework. *Cogent Engineering*, *7*(1). https://doi.org/10.1080/23311916.2020.1810862

Sarmast, Z., Shokouhyar, S., Ghanadpour, S. H., & Shokoohyar, S. (2023). Unravelling the potential of social media data analysis to improve the warranty service operation. *Industrial Management & Data Systems*, *123*(5), 1281–1309. https://doi.org/10.1108/IMDS-07-2022-0427

Singgalen, Y. A. (2023a). Analisis Perilaku Wisatawan Berdasarkan Data Ulasan di Website Tripadvisor Menggunakan CRISP-DM : Wisata Minat Khusus Pendakian Gunung Rinjani dan Gunung Bromo. *Journal of Computer System and Informatics (JoSYC)*, *4*(2), 326–338. https://doi.org/10.47065/josyc.v4i2.3042

Singgalen, Y. A. (2023b). Analisis Sentimen Pengunjung Pulau Komodo dan Pulau Rinca di Website Tripadvisor Berbasis CRISP-DM. *Journal of Information System Research (JOSH)*, *4*(2), 614–625. https://doi.org/10.47065/josh.v4i2.2999

Singgalen, Y. A. (2023c). Analisis Sentimen Wisatawan terhadap Kualitas Layanan Hotel dan Resort di Lombok Menggunakan SERVQUAL dan CRISP-DM. *Building of Informatics, Technology and Science (BITS)*, *4*(4), 1870–1882. https://doi.org/10.47065/bits.v4i4.3199

Singgalen, Y. A. (2023d). Analisis Sentimen Wisatawan terhadap Taman Nasional Bunaken dan Top 10 Hotel Rekomendasi Tripadvisor Menggunakan Algoritma SVM dan DT berbasis CRISP-DM. *Journal of Computer System and Informatics (JoSYC)*, *4*(2), 367–379. https://doi.org/10.47065/josyc.v4i2.3092

Singgalen, Y. A. (2023e). Penerapan CRISP-DM dalam Klasifikasi Sentimen dan Analisis Perilaku Pembelian Layanan Akomodasi Hotel Berbasis Algoritma Decision Tree (DT). *Jurnal Sistem Komputer Dan Informatika (JSON)*, *5*(2), 237–248. https://doi.org/10.30865/json.v5i2.7081

Singgalen, Y. A. (2023f). Penerapan Metode CRISP-DM dalam Klasifikasi Data Ulasan Pengunjung Destinasi Danau Toba Menggunakan Algoritma Naïve Bayes Classifier (NBC) dan Decision Tree (DT). *Jurnal Media Informatika Budidarma*, *7*(3), 1551–1562. https://doi.org/10.30865/mib.v7i3.6461

Singgalen, Y. A. (2023g). Sentiment classification of coral reef 101 content using decision tree algorithm through

CRISP-DM. *International Journal of Basic and Applied Science*, *12*(3), 121–130.

Srivastava, P. R., Sharma, D., & Kaur, I. (2022). Differential effects of online signals on sales performance of local brand clothing products. *Journal of Enterprise Information Management*, *35*(6), 1522–1547. https://doi.org/10.1108/JEIM-01-2020-0039

Suri, A., Huang, B., & Sénécal, S. (2023). This Product Seems Better Now: How Social Media Influencers' Opinions Impact Consumers' Post-failure Responses. *International Journal of Electronic Commerce*, *27*(3), 297–323. https://doi.org/10.1080/10864415.2023.2226898

van Atteveldt, W., van der Velden, M. A. C. G., & Boukes, M. (2021). The Validity of Sentiment Analysis: Comparing Manual Annotation, Crowd-Coding, Dictionary Approaches, and Machine Learning Algorithms. *Communication Methods and Measures*, *15*(2), 121–140. https://doi.org/10.1080/19312458.2020.1869198

Vanderkooi, D., Mashatan, A., & Turetken, O. (2023). Introducing technological disruption: how breaking media attention on corporate events impacts online sentiment. *Journal of Business Analytics*, *7*(2), 63–82. https://doi.org/10.1080/2573234X.2023.2274088

Wei Lun Lee, A., Ying Chung, S., Shee Tan, Y., Mun Ho Koh, S., Feng Lu, W., & Sze Choong Low, J. (2023). Enhancing the environmental sustainability of product through ecodesign: a systematic review. *Journal of Engineering Design*, *34*(10), 814–843. https://doi.org/10.1080/09544828.2023.2261094

Yerpude, S., & Rautela, S. (2022). Digitally driven new product development: an involved contemporary innovation case. *International Journal of Productivity and Performance Management*, *71*(5), 1953–1971. https://doi.org/10.1108/IJPPM-09-2019-0448

Zhu, S., Du, S., Nie, T., & Zhu, Y. (2023). Product Pricing and Design Strategy in Platform-Based Collaborative Innovation With Cognitive Bias. *International Journal of Electronic Commerce*, *27*(2), 236–269. https://doi.org/10.1080/10864415.2023.2184240