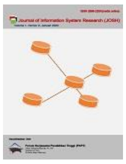

Journal of Information System Research (JOSH)
Volume 5, No. 3, April 2024, pp 841−852
ISSN 2686-228X (media online)
https://ejurnal.seminar-id.com/index.php/josh/
DOI 10.47065/josh.v5i3.5036

Sentiment and Topic Analysis of Aftermovie Piala Presiden eSports 2019 (Onic Esport)

Ruben Wiliam Setiawan, Abigail Rosandrine Kayla Putri Rahadi, 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: [1]ruben.202202530010@student.atmajaya.ac.id,[2] abigail.202202530014@student.atmajaya.ac.id, [3,*]yerik.afrianto@atmajaya.ac.id
Correspondence Author Email: yerik.afrianto@atmajaya.ac.id
Submitted: **23/03/2024;** Accepted: **30/04/2024;** Published: **30/04/2024**

Abstract−This research addresses the research problem of sentiment classification and topic analysis in the context of the Aftermovie Piala Presiden Esports 2019 (Onic Esport), employing the CRISP-DM methodology. Leveraging a dataset comprising 1830 posts, sentiment analysis was conducted on 191 posts using Vader and TextBlob algorithms, revealing the distribution of polarity values: 7.41% negative, 58.02% neutral, and 34.57% positive sentiments. Furthermore, the study evaluates the performance of classification algorithms, such as k-nearest Neighbors (k-NN), Support Vector Machine (SVM), and Decision Tree (DT), utilizing SMOTE. Notably, SVM demonstrated an accuracy of 90.34%, AUC of 0.995, precision of 99.91%, recall of 80.77%, and F-measure of 89.29%. Similarly, DT exhibited an accuracy of 94.83%, AUC of 0.950, precision of 92.12%, recall of 98.12%, and F-measure of 95.00%. K-NN displayed an accuracy of 92.83%, AUC of 0.972, precision of 98.01%, recall of 87.46%, and F-measure of 92.42%. The analysis also revealed frequently used words in the dataset, including 'onic' (127 occurrences), 'udil' (95 occurrences), and 'piala' (28 occurrences). This study sheds light on sentiment patterns and prevalent topics among esports enthusiasts, offering insights for content optimization and event management strategies.

Keywords: Sentiment; Topic; Piala Presiden; eSports; Onic

1. INTRODUCTION

Esports emerges as a captivating event for mobile gamers, proving advantageous for tourism. The burgeoning popularity of esports tournaments draws in diverse players, fostering a vibrant gaming community globally [1]–[3]. These events are pivotal platforms for showcasing skill and strategy in mobile gaming, thereby captivating the attention of enthusiasts and spectators alike [4]–[7]. Moreover, the substantial influx of attendees and participants to these tournaments stimulates local economies, bolstering tourism sectors in host cities [8]–[10]. Consequently, esports not only enriches the gaming landscape but also contributes significantly to the economic vitality of regions through increased tourism revenue [11].

Esports has emerged as a prestigious event that significantly influences purchasing behavior, particularly in digital consumption. The rise of competitive gaming has cultivated a dynamic market where players and fans avidly engage with digital content and merchandise [12]. These esports events are potent catalysts for driving consumer interest in gaming-related products and services, fueling a surge in digital consumption across various platforms [13]–[15]. Additionally, the allure of esports tournaments prompts brands to invest heavily in sponsorships and endorsements, further amplifying their influence on consumer purchasing decisions in the digital realm [16]–[18]. Consequently, esports is a pivotal force reshaping consumer behavior in the digital landscape, marking a paradigm shift towards heightened engagement and consumption within the gaming community [19].

This study proposes a novel approach to discern sentiments and analyze critical topics of interest among gamers and esports enthusiasts by examining user reviews within digital audio-visual content. By leveraging advanced sentiment analysis techniques and topic modeling algorithms, this research systematically extracts valuable insights from the vast pool of user-generated content in the form of reviews and comments [20]–[22]. Furthermore, this method offers a nuanced understanding of the gaming community's prevailing sentiments and focal points, facilitating informed decision-making processes for stakeholders in the esports industry [23]–[25]. Consequently, this research enhances our comprehension of the intricacies surrounding gamer preferences and perceptions, thereby contributing to advancing strategic initiatives and content development strategies within the esports domain [26].

The urgency of this research lies in its potential to address pressing challenges and opportunities within the burgeoning field of esports. With the gaming industry's rapid expansion and esports' growing influence on global entertainment and economy, there is a critical need to deepen our understanding of various facets within this domain [27]–[30]. By investigating key issues such as player behavior, audience engagement, and industry trends, this research provides valuable insights that inform strategic decision-making processes and drive sustainable growth in the esports ecosystem [31]. Therefore, the timely execution of this research is paramount in fostering innovation, enhancing competitiveness, and ensuring the long-term viability of the esports industry in an ever-evolving digital landscape [32].

Theoretical and practical implications of this research extend to purchasing behavior and digital consumption patterns within the tourism and esports event industries. This study contributes theoretical

Journal of Information System Research (JOSH)
Volume 5, No. 3, April 2024, pp 841−852
ISSN 2686-228X (media online)
https://ejurnal.seminar-id.com/index.php/josh/
DOI 10.47065/josh.v5i3.5036

frameworks that shed light on the intricate dynamics shaping purchasing decisions by elucidating the underlying drivers of consumer engagement and expenditure in these sectors. Furthermore, insights from this research offer actionable strategies for stakeholders, ranging from event organizers to marketers and policymakers, to optimize their approaches to the evolving needs and prefer problem-solvers. Consequently, this research catalyzes informed decision-making processes, fostering sustainable growth and innovation within the tourism and esports event domains, thereby enhancing their economic and cultural significance on a global scale.

The utilization of the CRISP-DM (Cross-Industry Standard Process for Data Mining) methodology in this research underscores its alignment with established best practices in data analysis and research methodology. While similar studies may have employed variations of this framework, the systematic and rigorous nature of CRISP-DM ensures a structured approach to data exploration, modeling, and interpretation, thereby enhancing the reliability and validity of the findings. However, it is imperative to acknowledge the limitations inherent in the CRISP-DM methodology, such as its reliance on predefined steps that may not adequately capture the complexity of specific research contexts or datasets. Nevertheless, by leveraging the CRISP-DM framework, this research endeavors to mitigate potential biases and enhance the robustness of its analytical processes, ultimately contributing to advancing knowledge in the field.

2. RESEARCH METHODOLOGY

2.1 Gap Analysis

The analysis of research gaps is essential for understanding the contributions to knowledge advancement, and this study undertakes a comprehensive examination of the evolution of research within the intersection of esports and tourism. By scrutinizing existing literature, this research aims to identify areas where substantial knowledge deficits persist, thereby pinpointing opportunities for further inquiry and innovation. Through this systematic analysis, the study endeavors to consolidate existing insights and illuminate avenues for future research endeavors that enrich our understanding of the dynamic relationship between esports and tourism. Consequently, the meticulous examination of research gaps catalyzes scholarly discourse and fosters new avenues of inquiry within this burgeoning field.

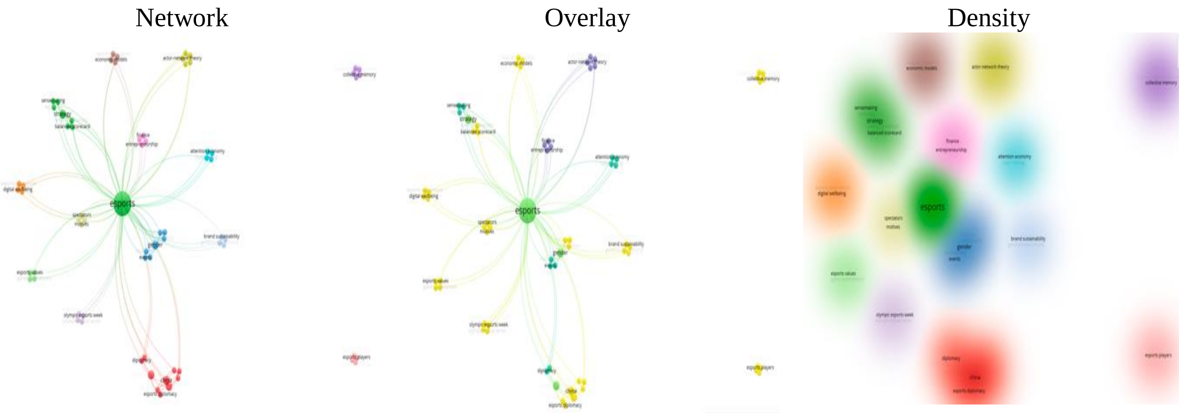

Figure 1. Network, Overlay, and Density Visualization using Vosviewer

Figure 1 shows the network, overlay, and density visualization using Vosviewer. Based on the analysis of research gaps, it is evident that studies focusing on sentiment and topic analysis for esports are crucial in understanding the behaviors of esports players. By delving into the sentiments expressed within user-generated content and identifying prevalent topics of discussion, this research gains valuable insights into the motivations, preferences, and attitudes driving player behavior in the esports ecosystem [33]. This underscores the significance of employing advanced analytical techniques to unravel the complexities of the gaming community, ultimately enhancing our understanding of the interplay between sentiment, topics, and player engagement in esports [34]–[37]. Consequently, prioritizing research in this domain holds great promise for advancing knowledge and informing strategic initiatives within the esports industry.

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

This study's sentiment and topic analysis methodology is the CRISP-DM (Cross-Industry Standard Process for Data Mining) framework. This structured approach encompasses distinct phases, including business understanding, data understanding, data preparation, modeling, evaluation, and deployment, ensuring a systematic and rigorous data analysis. By adhering to the CRISP-DM methodology, this research navigates the complexities inherent in sentiment and topic analysis, thus facilitating a comprehensive exploration of the gaming community's sentiments and interests within the esports domain. Consequently, using CRISP-DM is a robust foundation for deriving meaningful insights and informing strategic decision-making processes in esports research.

Journal of Information System Research (JOSH)
Volume 5, No. 3, April 2024, pp 841−852
ISSN 2686-228X (media online)
https://ejurnal.seminar-id.com/index.php/josh/
DOI 10.47065/josh.v5i3.5036

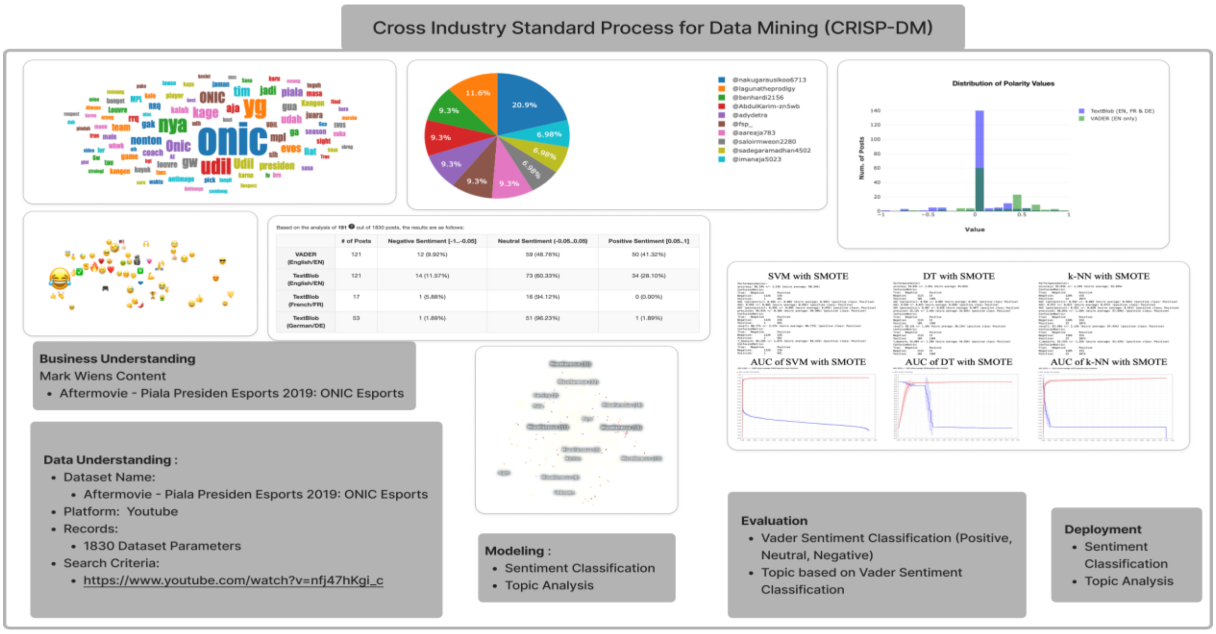

Figure 2. CRISP-DM Framework

Figure 2 shows the CRISP-DM framework. In the CRISP-DM framework, sentiment and topic analysis of video content is highly pertinent in discussing esports and tourism. Video content is a primary medium for communication and engagement in esports, offering rich insights into player sentiments and preferences. By systematically analyzing the sentiments expressed in these videos, this research discerns nuanced attitudes and emotions prevalent among gamers, thus informing strategic decision-making processes for stakeholders within the esports and tourism industries. Furthermore, identifying key topics discussed in video content related to esports and tourism enables a comprehensive understanding of the themes driving player engagement and audience interest. Consequently, integrating sentiment and topic analysis into the CRISP-DM framework facilitates a holistic exploration of the complex dynamics shaping the intersection of esports and tourism, ultimately contributing to informed strategies and initiatives for industry growth and development.

2.2.1 Business Understanding

During the business understanding phase, a comprehensive examination is conducted regarding the context of the analyzed videos, alongside the textual data extracted from reviews by both viewers and esports players. This initial step serves as the foundation for understanding the overarching objectives of the analysis and delineating the key variables and parameters to be considered throughout the subsequent phases of the CRISP-DM framework. By contextualizing the video content and integrating insights from textual reviews, this research gains a holistic understanding of the content landscape, facilitating informed decisions regarding the sentiment and topic analysis methodologies. Consequently, the meticulous examination conducted during the business understanding phase lays the groundwork for a robust and effective analytical process, ultimately enhancing the validity and reliability of the findings derived from the CRISP-DM framework.

Within the context of this research, the analyzed video titled "Aftermovie - Piala Presiden Esports 2019: ONIC Esports," sourced from the YouTube platform, emerges as a focal point. With a notable viewership of 2,864,378 and premiering on October 11, 2019, this video signifies a significant event within the esports domain. Moreover, its substantial engagement is evidenced by 1,830 comments, reflecting the diverse perspectives and reactions elicited from viewers and participants alike. This video serves as a rich data source for sentiment and topic analysis, offering valuable insights into the sentiments, preferences, and discussions prevalent within the esports community. Consequently, thoroughly examining this video content holds immense potential for elucidating key trends and dynamics shaping the intersection of esports and tourism, thereby contributing to advancing knowledge in this field.

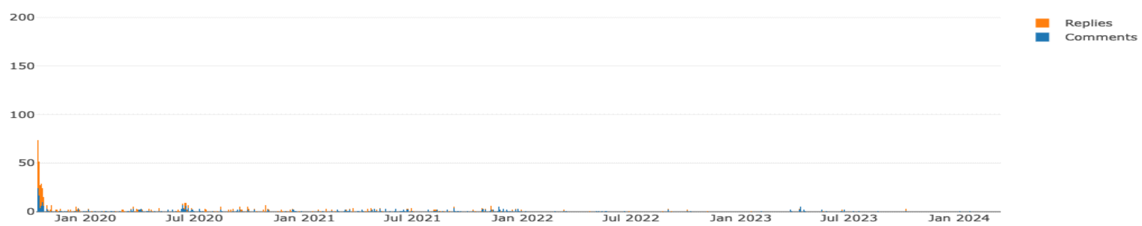

Figure 3. Post-per-day statistic of the Content (Communalytic)

Journal of Information System Research (JOSH)
Volume 5, No. 3, April 2024, pp 841−852
ISSN 2686-228X (media online)
https://ejurnal.seminar-id.com/index.php/josh/
DOI 10.47065/josh.v5i3.5036

Figure 3 shows the post-per-day statistic of the content. Based on the statistical data on post-per-day, it is discernible that there were varying levels of engagement with the analyzed video "Aftermovie - Piala Presiden Esports 2019: ONIC Esports" on different dates. Notably, on October 12, 2019, there were 84 comments and 111 replies, indicating a surge in interaction and discussion surrounding the video content. Conversely, on subsequent days, such as October 16th and 22nd, the engagement dwindled with only 16 and 14 comments, respectively. This fluctuation in engagement underscores the dynamic nature of audience interaction and highlights the importance of temporal considerations in analyzing viewer responses to video content. Consequently, this data provides valuable insights into the patterns of engagement and the factors influencing audience participation, thereby enriching our understanding of the dynamics within the esports community.

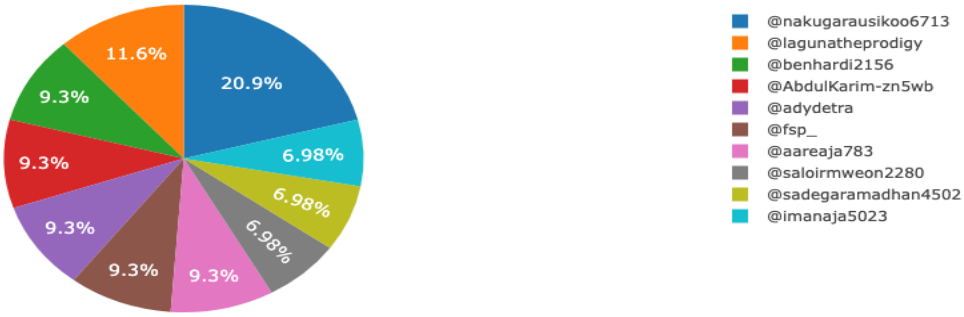

Figure 4. Top Ten Poster of the Content (Communalytic)

Figure 4 shows the top ten posters of the content. Based on the statistical data on the top ten posters, it is evident that @nakugarausikoo6713 emerged as the most prolific contributor with nine posts, followed by @lagunatheprodigy with five posts. The subsequent posters, including @benhardi2156, @AbdulKarim-zn5wb, @adydetra, @fsp_, @aareaja783, @saloirmweon2280, @sadegaramadhan4502, and @imanaja5023, each made 3 to 4 posts. This posting activity distribution underscores the analyzed community's varying engagement and contribution levels. Moreover, it offers valuable insights into the individuals actively participating and shaping discussions surrounding the analyzed video content. Consequently, this data aids in understanding the dynamics of user engagement and identifying key influencers within the esports community, thereby informing strategic approaches for content dissemination and community management.

Through the statistical data on post-per-day and the identification of the top ten posters, valuable insights into YouTube users' intentions and interests regarding the Esports President's Cup 2019 content are discernible. The fluctuations in engagement levels observed across different dates highlight the dynamic nature of user interest, suggesting varying degrees of enthusiasm or attention toward the analyzed content. Additionally, the distribution of posting activity among top contributors indicates individuals who are actively engaged with and potentially influential within the esports community. Consequently, combining these data sets provides a comprehensive understanding of user intentions and preferences, offering valuable guidance for content creators, marketers, and organizers in tailoring their strategies to engage with the target audience effectively.

2.2.2 Data Understanding

Identifying frequently used words within the review dataset takes precedence during the data understanding phase. This initial step involves systematically parsing through the textual data to discern recurring terms or phrases that hold significance within the context of the analyzed content. This research gains insights into prevalent themes, sentiments, and discussion topics among viewers and players by pinpointing these frequently used words. Consequently, this process lays the groundwork for subsequent analyses, guiding this research in uncovering patterns and trends that inform the overarching objectives of the study and contribute to a nuanced understanding of audience perceptions and behaviors.

Figure 5. Frequently Used Words (Communalytic)

Journal of Information System Research (JOSH)
Volume 5, No. 3, April 2024, pp 841−852
ISSN 2686-228X (media online)
https://ejurnal.seminar-id.com/index.php/josh/
DOI 10.47065/josh.v5i3.5036

Figure 5 shows the frequently used words in word clouds. Based on the results of identifying frequently used words, it becomes apparent that specific terms dominate the discourse surrounding the analyzed content. Key terms such as "Onic" (127 occurrences), "Udil" (59 occurrences), "evos" (29 occurrences), "mpl" (27 occurrences), and "piala" (28 occurrences) prominently feature among the most frequently used words, indicating their significance within the context of esports discussions. These terms likely represent notable players, teams, tournaments, and game-related concepts relevant to viewers and players alike. Moreover, the repetition of specific terms like "yg" (75 occurrences), "nya" (60 occurrences), "gw" (37 occurrences), and "aja" (29 occurrences) suggests the informal language commonly employed in online discussions within the esports community. Consequently, analyzing frequently used words offers valuable insights into the prevalent themes, topics, and linguistic patterns present within the dataset, contributing to a deeper understanding of audience engagement and preferences within the esports domain.

Figure 6. Emoji Clouds (Communalytic)

Figure 6 shows the emoji clouds using Communalytic. Based on the results of identifying emoji clouds, it is evident that certain emojis are frequently used within the discourse surrounding the analyzed content. Notably, the "□" emoji appears 32 times, followed by "□," "□," and "□" with ten occurrences each. These emojis, including "□," "□," and "□," among others, convey a range of emotions and reactions commonly expressed by users in online discussions, reflecting the dynamic and emotive nature of engagement within the esports community. Additionally, the presence of emojis such as "□," "□," and "□" suggests positive sentiments and expressions of support or appreciation prevalent among viewers and players. Consequently, analyzing emoji clouds offers valuable insights into the dataset's emotional tone, sentiment, and engagement patterns, enhancing our understanding of audience perceptions and interactions within the esports domain.

Popular terms and reviewer sentiments were identified by analyzing frequently used words and emojis, providing valuable insights into audience engagement and perceptions within the analyzed content. By examining the recurrence of specific words and symbols, this research discerns prevalent themes, topics, and emotional responses expressed by viewers and players. This comprehensive approach facilitates a nuanced understanding of audience preferences and interactions, enabling content creators, marketers, and organizers to tailor their strategies to effectively engage with the target audience. Consequently, integrating frequently used words and emojis offers a holistic perspective on audience sentiments and preferences, contributing to informed decision-making processes and content development strategies within the esports domain.

2.2.3 Modeling

During the modeling phase, the process involves sentiment classification using Vader and Textblob, followed by topic analysis based on sentiment classification results. This systematic approach entails utilizing advanced natural language processing techniques to classify the sentiment expressed in textual data, enabling this research to discern positive, negative, or neutral sentiments. Subsequently, the classified sentiments serve as a basis for conducting topic analysis, allowing this research to identify prevalent themes, subjects, and discussions within the dataset. By integrating sentiment classification and topic analysis, this research underlying sentiments driving specific topics and conversations within the analyzed content, thus facilitating nuanced insights into audience perceptions and preferences within the esports domain. Consequently, this methodological approach enhances the depth and breadth of analysis, enabling this research to uncover meaningful patterns and trends that inform strategic decision-making processes and content development strategies.

The algorithms employed in model testing include k-NN, DT, NBC, and SVM, both with and without utilizing SMOTE. The best-performing model, determined through rigorous evaluation and comparison, will be advanced for further discussion and analysis. This methodological approach ensures a comprehensive assessment of various machine learning algorithms and the impact of SMOTE, an oversampling technique, on model

Journal of Information System Research (JOSH)
Volume 5, No. 3, April 2024, pp 841−852
ISSN 2686-228X (media online)
https://ejurnal.seminar-id.com/index.php/josh/
DOI 10.47065/josh.v5i3.5036

performance. This research identifies the most effective approach for sentiment classification and topic analysis within the study context by systematically evaluating multiple algorithms and their variations. Consequently, selecting the optimal model will facilitate informed discussions and insights into audience sentiments and preferences within esports, contributing to advancing knowledge in this field.

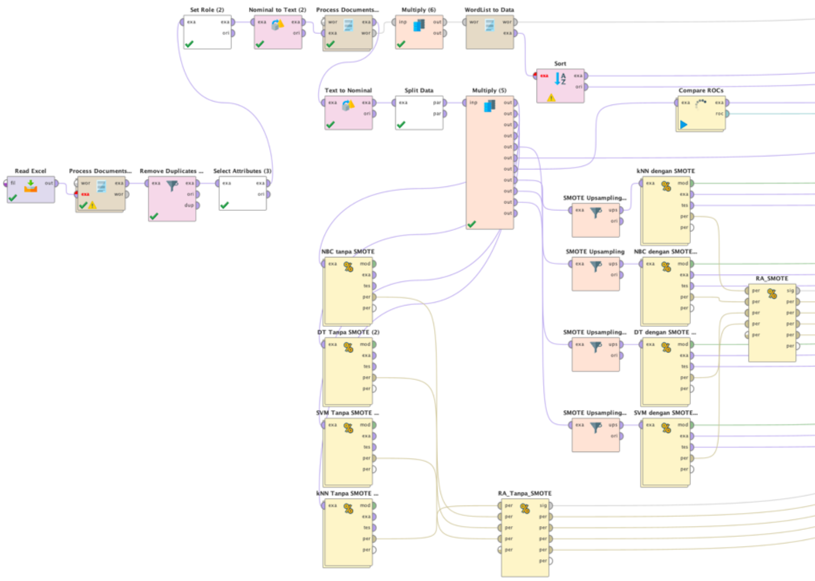

Figure 7. Modeling Process Using NBC, DT, k-NN, and SVM (Rapidminer)

Figure 7 shows the modeling process using Rapidminer. In the process of testing algorithm performance, the testing data is partitioned into 70% for testing and 30% for training, drawn from a total dataset comprising 1817 text entries in Indonesian and English. This systematic division ensures a robust evaluation of algorithm efficacy by utilizing a sizable portion of the dataset for training purposes while reserving a distinct subset for validation. Such a methodological approach enables this research to assess the generalization capability of the models across diverse linguistic contexts and effectively mitigate potential biases or overfitting issues. Consequently, this balanced allocation of data facilitates rigorous testing and validation procedures, ultimately enhancing the reliability and robustness of the conclusions drawn from the study.

Meanwhile, topic analysis employs the GloVe model, which utilizes global word embeddings to represent semantic relationships between words within a corpus. This approach leverages pre-trained word vectors to capture intricate semantic nuances and associations, facilitating a comprehensive exploration of topics and themes present within the analyzed textual data. By harnessing the power of distributed word representations, the GloVe model enables this research to uncover latent topics and discern underlying patterns that may not be readily apparent through traditional methods. Consequently, utilizing the GloVe model enriches the depth and granularity of topic analysis, enhancing our understanding of the complex dynamics and discussions within the dataset.

2.2.4 Evaluation

The evaluation of algorithm performance to serve as a sentiment analysis model by the dataset in this study considers multiple metrics, including accuracy, precision, recall, F-measure, and AUC (Area Under the Curve). These metrics collectively provide a comprehensive assessment of the model's effectiveness in accurately classifying sentiments expressed within the dataset. Accuracy measures the overall correctness of predictions, while precision quantifies the proportion of correctly predicted positive sentiments among all predicted positive instances. Recall evaluates the proportion of correctly identified positive sentiments among all actual positive instances, and the F-measure harmonizes precision and recall into a single score. Additionally, AUC assesses the model's ability to distinguish between positive and negative sentiments across various threshold values. By evaluating performance based on these metrics, this research informed decisions regarding the suitability and efficacy of sentiment analysis models for the specific requirements of the study dataset, thereby enhancing the reliability and validity of the research findings.

Meanwhile, the GloVe model is evaluated based on the sentiment classification results obtained from Vader (positive, neutral, negative), which are then visualized using Atlas Nomic. This approach allows for a comprehensive assessment of the GloVe model's performance in capturing sentiment nuances within the textual data. By leveraging the sentiment classifications generated by Vader, this research analyzes the alignment between the sentiment predictions made by the GloVe model and the ground truth sentiments, thus gauging the model's accuracy and efficacy. Moreover, the visualization of these sentiment classifications through Atlas Nomic facilitates a clear and intuitive understanding of the model's performance, aiding this research in identifying areas

Journal of Information System Research (JOSH)
Volume 5, No. 3, April 2024, pp 841−852
ISSN 2686-228X (media online)
https://ejurnal.seminar-id.com/index.php/josh/
DOI 10.47065/josh.v5i3.5036

for improvement and refinement. Consequently, this combined approach enhances the comprehensiveness and interpretability of the evaluation process, contributing to advancing sentiment analysis methodologies within the field.

2.2.5 Deployment

The deployment phase of the research findings on sentiment classification and topic analysis of the Aftermovie Piala Presiden Esports 2019 (Onic Esport) involves implementing and disseminating the developed models and insights derived from the study. This process includes integrating the sentiment classification and topic analysis methodologies into relevant industry applications or platforms to enhance decision-making processes and audience engagement strategies within the esports domain. Furthermore, disseminating research findings through academic publications, conferences, and industry partnerships ensures the wider adoption and utilization of the developed models and insights by stakeholders in the esports community. Consequently, the deployment phase is crucial in translating research outcomes into actionable solutions that advance sentiment analysis and topic analysis methodologies in esports and related domains.

3. RESULT AND DISCUSSION

The discussion of the research findings is bifurcated into two sections: the results of sentiment classification and topic analysis concerning the Aftermovie Piala Presiden Esports 2019 (Onic Esport). This structured approach allows for a comprehensive examination of the study outcomes, enabling this research to delve into the nuances of sentiment expressions and prevalent topics within the analyzed content. By segregating the discussion into these distinct components, this research provides in-depth insights into the sentiment dynamics and thematic trends observed in the Aftermovie, facilitating a nuanced understanding of audience perceptions and preferences within the esports community.

3.1 Sentiment Classification

Based on the sentiment analysis results conducted using CommAnalytic on 191 posts out of 1830 posts, the performance of Vader and TextBlob (English) in terms of the distribution of polarity values is evident. The analysis reveals that six posts (7.41%) exhibit negative sentiments, characterized by polarity scores equal to or less than -0.05. In comparison, 47 posts (58.02%) are classified as neutral sentiments, with polarity scores between -0.05 and 0.05. Moreover, 28 posts (34.57%) demonstrate positive sentiments, with polarity scores equal to or greater than 0.05. This detailed breakdown provides valuable insights into the sentiment dynamics present within the dataset, highlighting the varying degrees of positivity, neutrality, and negativity expressed across the analyzed posts. Consequently, this analysis underscores the efficacy of Vader and TextBlob in accurately capturing sentiment nuances within the dataset, thereby enhancing our understanding of audience sentiments and preferences within the study context.

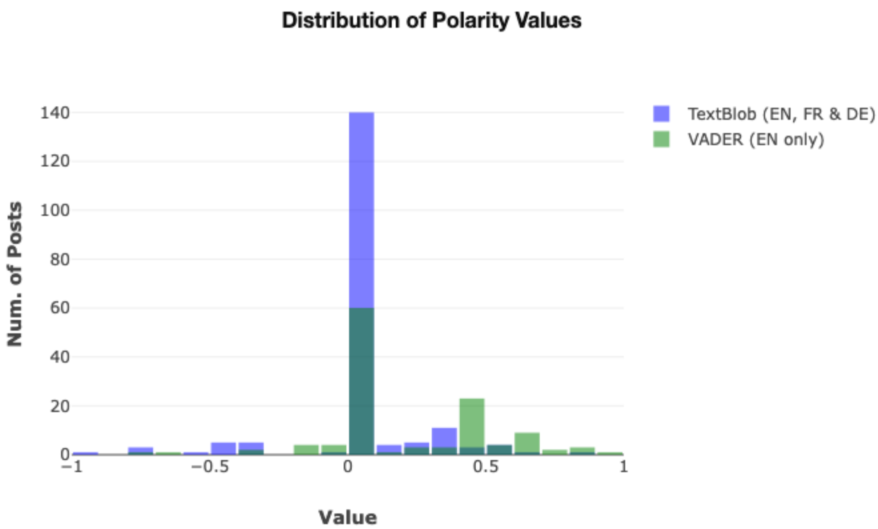

Figure 7. Distribution of Polarity Values (Communalytic)

Figure 7 shows the distribution of polarity values using Communalytic. The data processing follows a systematic approach with the following stages: Before running the analysis, the Sentiment Analysis Module excludes elements such as @usernames, #hashtags, $cashtags, and URLs to ensure the focus remains on the textual content. Subsequently, posts in English undergo analysis using both VADER and TextBlob sentiment analysis

Journal of Information System Research (JOSH)
Volume 5, No. 3, April 2024, pp 841−852
ISSN 2686-228X (media online)
https://ejurnal.seminar-id.com/index.php/josh/
DOI 10.47065/josh.v5i3.5036

tools, allowing for a comprehensive assessment of sentiment dynamics. Conversely, posts written in French or German are exclusively analyzed using TextBlob, catering to language-specific nuances. However, posts in languages other than English, French, or German are excluded from analysis. This structured data processing methodology ensures a thorough and accurate evaluation of sentiment across diverse linguistic contexts, thereby enhancing the reliability and validity of the research findings.

Based on the findings of the pilot testing conducted on 191 posts, further testing was carried out on a larger scale, encompassing 1817 posts from the review data. This expanded testing phase employed a structured approach with a data partitioning setting of 70% for testing and 30% for training within the RapidMiner application. Such a methodological framework ensures a robust evaluation of the developed models and algorithms, allowing for comprehensive testing across diverse data samples. This research enhances the models' accuracy and generalization capability by systematically partitioning the data and utilizing a significant portion for training, yielding more reliable and insightful results. Consequently, this systematic testing approach contributes to advancing sentiment analysis and topic analysis methodologies within the research domain.

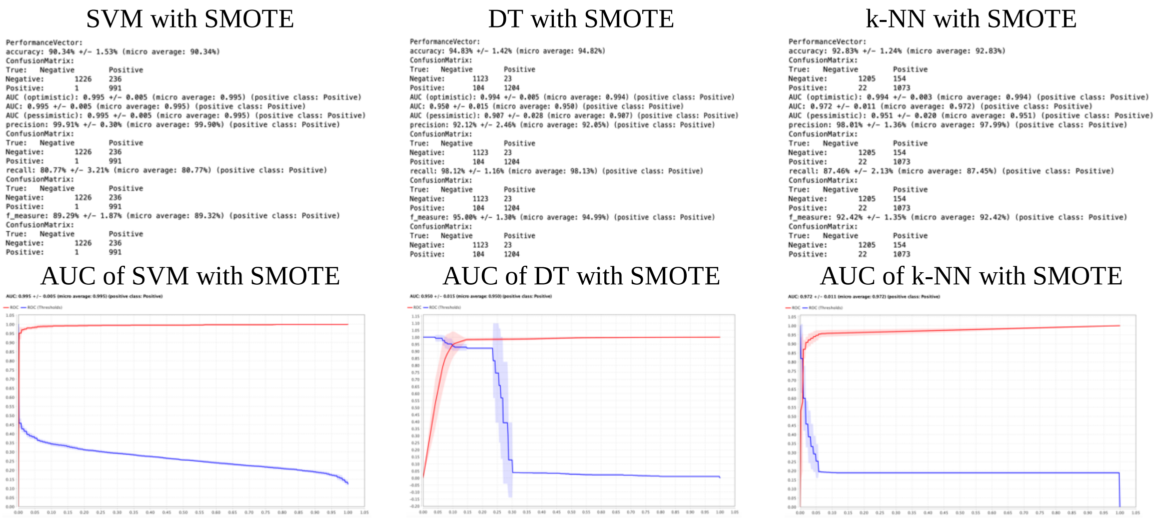

Figure 8. SVM, DT, and k-NN Model Performance (Rapidminer)

Figure 8 shows the best performance of the SVM, DT, and k-NN models in sentiment classification. Based on the performance evaluation results, it is evident that SVM utilizing SMOTE demonstrates commendable performance metrics as follows: accuracy of 90.34%, AUC (Area Under the Curve) of 0.995, precision of 99.91%, recall of 80.77%, and F-measure of 89.29%. In addition, the Decision Tree (DT) model utilizing SMOTE demonstrates commendable performance metrics as follows: an accuracy of 94.83%, an AUC (Area Under the Curve) of 0.950, a precision of 92.12%, recall of 98.12%, an F-measure of 95.00%. Otherwise, the k-Nearest Neighbors (k-NN) model utilizing SMOTE demonstrates commendable performance metrics as follows: an accuracy of 92.83%, an AUC (Area Under the Curve) of 0.972, precision of 98.01%, recall of 87.46%, and F-measure of 92.42%. The results of sentiment classification indicate that the k-nearest Neighbors (k-NN), Support Vector Machine (SVM), and Decision Tree (DT) algorithms utilizing the SMOTE operator exhibit commendable performance in classification tasks.

These algorithms demonstrate robust capabilities in accurately categorizing sentiment classes within the analyzed dataset. The effectiveness of k-NN, SVM, and DT models with SMOTE is evident from their high accuracy, precision, recall, and F-measure metrics, showcasing their ability to effectively discern positive, negative, and neutral sentiments. This performance underscores the suitability and efficacy of these algorithms for sentiment classification tasks, thereby contributing to the advancement of sentiment analysis methodologies within the research domain. Modeling topics using GloVe embeddings facilitated through sentiment classification. By first categorizing text data into sentiment classes, such as positive, negative, or neutral, sentiment analysis provides a structured foundation for further analysis. GloVe embeddings capture semantic relationships between words and generate vector representations by leveraging this sentiment-labeled data. These representations enable the modeling of topics within the corpus by identifying clusters of semantically related words or phrases. Thus, sentiment classification is a precursor to topic modeling, offering valuable insights into the underlying sentiments expressed within the text data and facilitating more nuanced analysis.

3.2 Topic Analysis

Topic analysis was conducted to classify sentiments into positive, neutral, and negative categories, providing a comprehensive understanding of textual data's underlying themes and patterns. By categorizing sentiments, this research identifies prevalent topics and analyzes their distribution across different sentiment classes, thereby gaining insights into the nuanced perceptions and opinions of the text. Furthermore, these topics are effectively

Journal of Information System Research (JOSH)
Volume 5, No. 3, April 2024, pp 841−852
ISSN 2686-228X (media online)
https://ejurnal.seminar-id.com/index.php/josh/
DOI 10.47065/josh.v5i3.5036

visualized through atlas nomic, enabling this research to explore and interpret the relationships between topics and sentiments in a structured and intuitive manner. As such, sentiment-based topic analysis and visualization tools like Atlas Nomic offer a powerful approach to uncovering meaningful insights from textual data, facilitating a more profound understanding and interpretation of complex information.

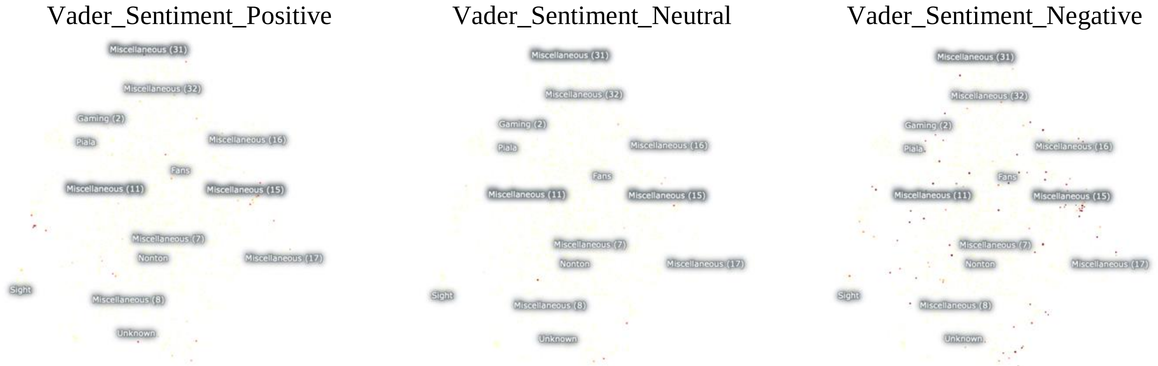

Figure 9. Topic Analysis based on Vader Sentiment (Atlas Nomic)

Figure 9 shows the topic visualization using Atlas Nomic based on positive, neutral, and negative Vader sentiments. Based on the results of the topic analysis, it becomes possible to map out the most frequently discussed topics based on video content, thereby providing valuable insights for event organizers and content creators to enhance performance and effectiveness, both in terms of events and digital content. By identifying prevalent themes and subjects within video content discourse, stakeholders tailor their strategies to better resonate with their target audience and meet their expectations. This process enables a more informed approach to content creation and event planning, improving engagement, satisfaction, and overall success. As such, leveraging the findings of topic analysis serves as a strategic tool for optimizing performance and achieving desired outcomes in both event management and digital content creation endeavors.

3.3 Discussion

The relationship between tourism and esports is mutually beneficial, as hosting international-scale events significantly boosts visitor arrivals to the host country. With their global appeal and extensive viewership, Esports tournaments attract a diverse audience worldwide, generating heightened interest and attention toward the host destination [38]. This influx of visitors contributes to the local tourism industry's economic growth. It enhances the country's international visibility and reputation as a vibrant and dynamic destination for esports enthusiasts and travelers. Consequently, the synergy between tourism and esports catalyzes fostering cross-cultural exchange, economic development, and global connectivity, thereby underscoring the symbiotic nature of their relationship.

The digital content in the form of the after-movie of the 2019 Presidential Cup Esports serves as both a catalyst and a stimulus for the behavior of gamers, warranting the identification and analysis of purchasing behavior related to products and services. As a visually compelling representation of the esports event, the after-movie captures the excitement and energy of the tournament, effectively engaging and influencing the gaming community [39], [40]. This engagement extends beyond entertainment, prompting viewers to consider their preferences, interests, and potential investments in gaming-related products and services [33], [41], [42]. Therefore, analyzing the impact of such digital content on consumer behavior is essential for understanding the dynamics of the gaming market and devising targeted strategies to cater to the needs and preferences of gamers effectively.

The findings of this study indicate that video content reviewers are both gamers and enthusiasts of esports events, highlighting a market that should not be underestimated. These reviewers demonstrate their active involvement and interest in the gaming community and esports scene by engaging with and providing feedback on esports-related content [43]–[46]. Their dual role as players and fans underscores the significance of their perspectives and opinions in shaping the perception and reception of esports content [37], [47]–[49]. Thus, acknowledging the influence and significance of content reviewers as part of the gaming ecosystem is essential for understanding consumer behavior and devising effective strategies to cater to the diverse needs and preferences within the esports market.

Through the outcomes of sentiment analysis and topic analysis, discerning the preferences and behaviors of gamers intrigued by esports becomes feasible. These analytical processes delve into the sentiments expressed within gaming communities and dissect the prevalent topics that resonate among esports enthusiasts [50]. A comprehensive understanding of gamers' inclinations and tendencies towards esports emerges by scrutinizing sentiment patterns and topical trends. This insight is invaluable for stakeholders in the gaming industry, enabling them to tailor their offerings, events, and marketing strategies to effectively engage and cater to the desires of the esports audience.

4. CONCLUSION

In conclusion, the research sheds light on the intricate dynamics of the esports landscape, uncovering valuable insights into gamers' sentiments, preferences, and behaviors. The study reveals the nuanced nuances within gaming communities through comprehensive sentiment analysis and topic analysis, providing a deeper understanding of what resonates with esports enthusiasts. Based on the performance evaluation results, it is evident that SVM utilizing SMOTE demonstrates commendable performance metrics as follows: accuracy of 90.34%, AUC (Area Under the Curve) of 0.995, precision of 99.91%, recall of 80.77%, and F-measure of 89.29%. In addition, the Decision Tree (DT) model utilizing SMOTE demonstrates commendable performance metrics as follows: an accuracy of 94.83%, an AUC (Area Under the Curve) of 0.950, a precision of 92.12%, recall of 98.12%, an F-measure of 95.00%. Otherwise, the k-Nearest Neighbors (k-NN) model utilizing SMOTE demonstrates commendable performance metrics as follows: an accuracy of 92.83%, an AUC (Area Under the Curve) of 0.972, precision of 98.01%, recall of 87.46%, and F-measure of 92.42%. The results of sentiment classification indicate that the k-nearest Neighbors (k-NN), Support Vector Machine (SVM), and Decision Tree (DT) algorithms utilizing the SMOTE operator exhibit commendable performance in classification tasks. Through these findings, stakeholders in the gaming industry enhance engagement, foster community connection, and drive growth within the rapidly evolving realm of esports.

ACKNOWLEDGMENT

We sincerely thank the Tourism Study Program, Faculty of Business Administration and Communication, Atma Jaya Catholic University of Indonesia, for their unwavering support and invaluable contributions to this research endeavor. Their guidance and resources have been instrumental in shaping the trajectory of this study, enabling us to delve deeper into the realms of sentiment analysis, social network dynamics, and toxicity assessment within the esports landscape. This collaborative effort underscores the commitment of the academic community to fostering innovation and knowledge dissemination.

REFERENCES

[1] A. Calapez, T. Ribeiro, V. Almeida, and V. Pedragosa, “Esports fan identity toward sponsor–sponsee relationship: an understanding of the role-based identity,” *Int. J. Sport. Mark. Spons.*, vol. 25, no. 1, pp. 42–66, Jan. 2024, doi: 10.1108/IJSMS-02-2023-0030.

[2] A. Z. Abbasi, M. Asif, A. Shamim, D. H. Ting, and R. A. Rather, “Engagement and consumption behavior of eSports gamers,” *Spanish J. Mark. - ESIC*, vol. 27, no. 2, pp. 261–282, Jan. 2023, doi: 10.1108/SJME-07-2022-0161.

[3] T. Ribeiro, V. Almeida, A. Calapez, H. Matsuoka, and R. Yamashita, “Esports and Olympic Games: a cross-cultural exploration of the player support behaviour towards the Olympics,” *Int. J. Sport. Mark. Spons.*, vol. 24, no. 4, pp. 700–721, Jan. 2023, doi: 10.1108/IJSMS-12-2022-0215.

[4] M. de Zoeten and T. Könecke, “Esports, video gaming and their fuzziness: a conceptualization and categorization,” *Sport. Bus. Manag. An Int. J.*, vol. 14, no. 1, pp. 99–116, Jan. 2024, doi: 10.1108/SBM-10-2022-0097.

[5] X. Wang, X. You, Y. Xu, and J. Zheng, “Online posting intention: do the social communication and brand equity of esports matter?,” *Int. J. Sport. Mark. Spons.*, vol. 25, no. 1, pp. 161–187, Jan. 2024, doi: 10.1108/IJSMS-09-2023-0189.

[6] B. D. A. Freitas, R. S. Contreras-Espinosa, and P. Á. P. Correia, “The importance of incorporating relevant-added value into esports sponsorships,” *J. Contemp. Mark. Sci.*, vol. 4, no. 3, pp. 397–423, Jan. 2021, doi: 10.1108/JCMARS-03-2021-0009.

[7] A. Z. Abbasi, M. Asif, L. D. Hollebeek, J. U. Islam, D. H. Ting, and U. Rehman, “The effects of consumer esports videogame engagement on consumption behaviors,” *J. Prod. Brand Manag.*, vol. 30, no. 8, pp. 1194–1211, Jan. 2021, doi: 10.1108/JPBM-04-2020-2839.

[8] K. Chang, J.-P. Uhm, S. Kim, and H.-W. Lee, “Paradoxical relationship between esports toxicity and toxicity tolerance: moderated mediation by gender and positive reappraisal coping,” *Int. J. Sport. Mark. Spons.*, vol. 24, no. 4, pp. 737–752, Jan. 2023, doi: 10.1108/IJSMS-11-2022-0197.

[9] J. Yadav, M. Misra, N. P. Rana, K. Singh, and S. Goundar, “Netizens’ behavior towards a blockchain-based esports framework: a TPB and machine learning integrated approach,” *Int. J. Sport. Mark. Spons.*, vol. 23, no. 4, pp. 665–683, Jan. 2022, doi: 10.1108/IJSMS-06-2021-0130.

[10] D. Gupta, H. Sharma, and M. Gupta, “Doping as a barrier in universal acceptance of esports,” *Int. J. Sport. Mark. Spons.*, vol. 23, no. 3, pp. 645–664, Jan. 2022, doi: 10.1108/IJSMS-05-2021-0105.

[11] S. Murray, J. Birt, and S. Blakemore, “eSports diplomacy: towards a sustainable ‘gold rush,’” *Sport Soc.*, vol. 25, no. 8, pp. 1419–1437, 2022, doi: 10.1080/17430437.2020.1826437.

[12] K. J. Mechelin and B. Liu-Lastres, “Exploring esport players’ motivation, experiences, and well-being,” *Leis. Stud.*, vol. 00, no. 00, pp. 1–15, 2023, doi: 10.1080/02614367.2023.2280043.

[13] M. Huettermann, G. T. Trail, A. D. Pizzo, and V. Stallone, “Esports Sponsorship: An Empirical Examination of Esports Consumers’ Perceptions of Non-Endemic Sponsors,” *J. Glob. Sport Manag.*, vol. 8, no. 2, pp. 524–549, 2023, doi: 10.1080/24704067.2020.1846906.

[14] G. Lam, O. Wai, K. Wong, and G. Lam, “Esports development ecosystem in Hong Kong : an application of a systems thinking literacy approach a systems thinking literacy approach,” *Leis. Stud.*, vol. 00, no. 00, pp. 1–18, 2024, doi: 10.1080/02614367.2024.2328086.

[15] A. Roth, R. E. Kunz, and C. Kolo, “The players’ perspective of value co-creation in esports service ecosystems,” *J. Media*

Journal of Information System Research (JOSH)
Volume 5, No. 3, April 2024, pp 841−852
ISSN 2686-228X (media online)
https://ejurnal.seminar-id.com/index.php/josh/
DOI 10.47065/josh.v5i3.5036

Bus. Stud., vol. 00, no. 00, pp. 1–25, 2023, doi: 10.1080/16522354.2023.2225299.
[16] A. D. Pizzo, G. J. Jones, B. J. Baker, D. C. Funk, and T. Kunkel, "Sensemaking of novelty: the dynamic nature of integrating esports within a traditional sport organization," *Sport Manag. Rev.*, vol. 25, no. 3, pp. 383–405, 2022, doi: 10.1080/14413523.2021.1935609.
[17] H. Abdolmaleki, A. D. Pizzo, B. J. Baker, A. Mahmoudi, and E. A. Ghahfarokhi, "Esports in Emerging Markets: A Balanced Scorecard Approach to LAN Gaming Centers in Iran," *J. Glob. Sport Manag.*, vol. 0, no. 0, pp. 1–21, 2023, doi: 10.1080/24704067.2023.2261444.
[18] Z. Lin and Y. Zhao, "(Dis)Assembling eSports: Material Actors and Relational Networks in the Chinese eSports Industry," *Crit. Arts*, vol. 35, no. 5–6, pp. 210–223, 2021, doi: 10.1080/02560046.2021.1986731.
[19] B. McCauley, K. Tierney, and D. Tokbaeva, "Shaping a Regional Offline eSports Market: Understanding How Jönköping, the 'City of DreamHack', Takes URL to IRL," *JMM Int. J. Media Manag.*, vol. 22, no. 1, pp. 30–48, 2020, doi: 10.1080/14241277.2020.1731513.
[20] Y. Zhao, Q. Li, and Z. Lin, "Toward cultural and creative industry: Chinese eSports through a business ecosystem lens," *J. Cult. Econ.*, vol. 16, no. 2, pp. 260–276, 2023, doi: 10.1080/17530350.2022.2159494.
[21] A. D. Pizzo, S. Na, D. Kim, K. Alexandris, and M. Hyun, "Esports gender diversity: A leisure constraints perspective," *J. Leis. Res.*, vol. 54, no. 5, pp. 602–623, 2023, doi: 10.1080/00222216.2023.2193186.
[22] W. Wang and J. Fan, "Topic mining of real-time discussions: what catches the attention of live-streaming esports viewers?," *Eur. Sport Manag. Q.*, pp. 1–22, 2022, doi: 10.1080/16184742.2022.2117397.
[23] Z. Belyaeva, A. Petrosyan, and S. M. R. Shams, "Stakeholder data analysis in the video gaming industry: implications for regional development," *EuroMed J. Bus.*, vol. 17, no. 3, pp. 333–349, Jan. 2022, doi: 10.1108/EMJB-10-2021-0150.
[24] J. P. Simon, "The production and distribution of digital content in China. An historical account of the role of internet companies and videogames," *Digit. Policy, Regul. Gov.*, vol. 23, no. 2, pp. 190–209, Jan. 2021, doi: 10.1108/DPRG-10-2020-0138.
[25] J. García and C. Murillo, "Sports video games participation: what can we learn for esports?," *Sport. Bus. Manag. An Int. J.*, vol. 10, no. 2, pp. 169–185, Jan. 2020, doi: 10.1108/SBM-01-2019-0006.
[26] M. Sjöblom, J. Macey, and J. Hamari, "Digital athletics in analogue stadiums," *Internet Res.*, vol. 30, no. 3, pp. 713–735, Jan. 2020, doi: 10.1108/INTR-07-2018-0304.
[27] D. Micallef, L. Parker, L. Brennan, B. Schivinski, and M. Jackson, "Emerging adult gamers and their diet – a socio-ecological approach to improve health behaviour," *J. Soc. Mark.*, vol. 14, no. 1, pp. 95–113, Jan. 2024, doi: 10.1108/JSOCM-03-2023-0066.
[28] X. M. S. Chen, L. Schuster, and E. Luck, "The well-being outcomes of multi-actor inter-organisational value co-creation and co-destruction within a service ecosystem," *J. Serv. Mark.*, vol. 37, no. 5, pp. 606–619, Jan. 2023, doi: 10.1108/JSM-03-2022-0082.
[29] A. Tjønndal, "'What's next? Calling beer-drinking a sport?!': virtual resistance to considering eSport as sport," *Sport. Bus. Manag. An Int. J.*, vol. 11, no. 1, pp. 72–88, Jan. 2021, doi: 10.1108/SBM-10-2019-0085.
[30] P. Parshakov, S. Paklina, D. Coates, and A. Chadov, "Does video games' popularity affect unemployment rate? Evidence from macro-level analysis," *J. Econ. Stud.*, vol. 48, no. 4, pp. 817–835, Jan. 2021, doi: 10.1108/JES-07-2019-0339.
[31] R. E. Kunz, A. Roth, and J. P. Santomier, "A perspective on value co-creation processes in eSports service ecosystems," *Sport. Bus. Manag. An Int. J.*, vol. 12, no. 1, pp. 29–53, Jan. 2022, doi: 10.1108/SBM-03-2021-0039.
[32] K. Lehnert, A. Walz, and R. Christianson, "The booming eSports market: a field day for fans," *J. Bus. Strategy*, vol. 43, no. 2, pp. 122–128, Jan. 2022, doi: 10.1108/JBS-07-2020-0159.
[33] W. W. Jang, K. K. Byon, T. A. Baker III, and Y. Tsuji, "Mediating effect of esports content live streaming in the relationship between esports recreational gameplay and esports event broadcast," *Sport. Bus. Manag. An Int. J.*, vol. 11, no. 1, pp. 89–108, Jan. 2021, doi: 10.1108/SBM-10-2019-0087.
[34] "Ways to influence consumption behaviors," *Strateg. Dir.*, vol. 37, no. 4, pp. 25–27, Jan. 2021, doi: 10.1108/SD-01-2021-0007.
[35] W. Hattingh, L. Van den Berg, and A. Bevan-Dye, "The 'why' behind generation Y amateur gamers' ongoing eSports gameplay intentions," *Int. J. Sport. Mark. Spons.*, vol. 25, no. 1, pp. 67–87, Jan. 2024, doi: 10.1108/IJSMS-04-2023-0064.
[36] D. L. Painter and B. Sahm, "Framing esports' JEDI issues: a case study in media irresponsibility," *Corp. Commun. An Int. J.*, vol. 28, no. 3, pp. 506–521, Jan. 2023, doi: 10.1108/CCIJ-06-2022-0068.
[37] A. Z. Abbasi, "Specifying, estimating and validating consumer eSports engagement composite model: a composite confirmatory approach," *EuroMed J. Bus.*, vol. 18, no. 3, pp. 452–466, Jan. 2023, doi: 10.1108/EMJB-04-2022-0068.
[38] H. Kohls, J. L. Hiler, and L. A. Cook, "Why do we twitch? Vicarious consumption in video-game livestreaming," *J. Consum. Mark.*, vol. 40, no. 6, pp. 639–650, Jan. 2023, doi: 10.1108/JCM-03-2020-3727.
[39] S. Anwar, S. Noor, and M. Asrar-ul-Haq, "Health risk behaviors among university students: the role of outdoor sports and electronic sports," *J. Humanit. Appl. Soc. Sci.*, vol. 5, no. 1, pp. 20–34, Jan. 2023, doi: 10.1108/JHASS-06-2021-0104.
[40] J. Yadav, M. Misra, N. P. Rana, and K. Singh, "Exploring the synergy between nano-influencers and sports community: behavior mapping through machine learning," *Inf. Technol. People*, vol. 35, no. 7, pp. 1829–1854, Jan. 2022, doi: 10.1108/ITP-03-2021-0219.
[41] J. Rietz and K. Hallmann, "A systematic review on spectator behavior in esports: why do people watch?," *Int. J. Sport. Mark. Spons.*, vol. 24, no. 1, pp. 38–55, Jan. 2023, doi: 10.1108/IJSMS-12-2021-0241.
[42] Y. Hwang, N. M. Watanabe, and M. Nagel, "Effects of in-game brand congruity on esports consumers' implicit and explicit memory," *Int. J. Sport. Mark. Spons.*, vol. 24, no. 1, pp. 108–128, Jan. 2023, doi: 10.1108/IJSMS-03-2022-0085.
[43] K. Chang, H.-W. Lee, and G. Bennett, "Generation Z and Millennial esports gamers' perceived toxicity risk and game brand loyalty: moderated mediation effects of brand identification, self-efficacy, and support from game brand," *Young Consum.*, vol. 25, no. 1, pp. 28–44, Jan. 2024, doi: 10.1108/YC-11-2022-1635.
[44] S.-Y. Chou, W. (William) Jang, S. C. Ma, C.-H. Chang, and K. K. Byon, "Is mobile gaming a new pillar of esports?

Volume 5, No. 3, April 2024, pp 841−852
ISSN 2686-228X (media online)
https://ejurnal.seminar-id.com/index.php/josh/
DOI 10.47065/josh.v5i3.5036

Exploring players' in-game purchases in PC and mobile platforms by using flow and clutch," *Int. J. Sport. Mark. Spons.*, vol. 24, no. 2, pp. 311–332, Jan. 2023, doi: 10.1108/IJSMS-07-2022-0134.

[45] W. (William) Jang, K. K. Byon, A. Williams, and P. M. Pedersen, "Augmenting the formation of esports gameplay intention: interaction effects of genre and gender," *Sport. Bus. Manag. An Int. J.*, vol. 11, no. 5, pp. 620–646, Jan. 2021, doi: 10.1108/SBM-04-2021-0049.

[46] S. J. Kelly and D. Van der Leij, "A new frontier: alcohol sponsorship activation through esports," *Mark. Intell. Plan.*, vol. 39, no. 4, pp. 533–558, Jan. 2021, doi: 10.1108/MIP-03-2020-0101.

[47] D. Ock and Y. Hwang, "The impacts of social needs and game playing type on esports consumers' positive eWOM sharing: applying the FIRO theory," *Int. J. Sport. Mark. Spons.*, vol. 24, no. 5, pp. 967–984, Jan. 2023, doi: 10.1108/IJSMS-02-2023-0022.

[48] E. Lettieri and C. Orsenigo, "Predicting soccer consumption: do eSports matter? Empirical insights from a machine learning approach," *Sport. Bus. Manag. An Int. J.*, vol. 10, no. 5, pp. 523–544, Jan. 2020, doi: 10.1108/SBM-10-2019-0093.

[49] R. Rogers, L. Farquhar, and J. Mummert, "Audience response to endemic and non-endemic sponsors of esports events," *Int. J. Sport. Mark. Spons.*, vol. 21, no. 3, pp. 561–576, Jan. 2020, doi: 10.1108/IJSMS-09-2019-0107.

[50] W. Wang, "Returning for skill or popularity? The demand for esports match replay," *Int. J. Sport. Mark. Spons.*, vol. 24, no. 2, pp. 295–310, Jan. 2023, doi: 10.1108/IJSMS-07-2022-0146.