Title	eスポーツゲームレビューのための効率的な観点感情分析フレームワーク
Author(s)	余, 洋
Citation	
Issue Date	2024-03
Type	Thesis or Dissertation
Text version	ETD
URL	http://hdl.handle.net/10119/19052
Rights	
Description	Supervisor:HUYNH, Van Nam, 先端科学技術研究科, 博士

Doctoral Dissertation

An Efficient Aspect-Based Sentiment Analysis Framework for Esports Game Reviews

YU Yang

Supervisor: HUYNH, Nam Van

Graduate School of Advanced Science and Technology
Japan Advanced Institute of Science and Technology
[Knowledge Science]
March, 2024

Abstract

This dissertation introduces an innovative framework for analyzing player feedback in esports games, with a focus on advanced data analysis to understand player emotions and game dynamics. The study begins with a comprehensive background, theoretical models, and empirical research on the proposed framework, introducing key concepts and terminology related to esports and player feedback analysis.

Initially, the research presents a thorough data analysis framework aimed at addressing the complex issues in esports player feedback. Utilizing a large dataset of approximately eight million reviews from major esports games on Steam and Google Play, including titles like PUBG, Dota2, CS:GO, and PUBG Mobile, the study offers a comprehensive analysis of player feedback. We enhance the topic modeling and sentiment analysis in this framework with the power of Transformer architecture, significantly improving the accuracy in interpreting player emotions and game dynamics. This method provides a novel approach to player feedback analysis, consistent with the statistical interpretation used in traditional data analysis.

The research also involves comparative analyses with existing models using popular evaluation methods in machine learning. The experimental results reveal that game optimization, server connectivity, anti-cheat mechanisms, and game updates are the top priorities for esports players currently. Generally, the insights not only demonstrate the ability of the enhanced topic modeling to reveal themes and sentiment analysis to uncover player emotions within the noisy feedback but also further illustrate the framework's completeness and the indispensable nature of each of its components. They are crucial for identifying common issues that resonate across different player groups, and invaluable for strategizing around game updates, community engagement, and player-centric approaches in game development. The adaptability and scalability of the framework make it an essential tool for the success of esports games.

Finally, the dissertation lays the foundation for future esports analytics research. It emphasizes the importance of advanced and detailed analytical tools in the evolving esports industry and their role in strengthening the symbiotic relationship between game developers and player communities. The research has been rigorously tested on various benchmarks, outperforming existing analytical models in efficiency, scalability, and depth of analysis. In summary, this dissertation provides a comprehensive and effective tool for analyzing player feedback in esports games, making a significant contribution to the field of esports analytics.

Keywords: Esports Gaming, Novice Player Experience, Player Feedback Analysis, Topic Modeling, Sentiment Analysis

Acknowledgments

I would like to express my sincere gratitude to all those who have played a crucial role in my academic journey and the successful completion of my PhD dissertation. This phase of my education has been immensely significant, and I am filled with deep appreciation for the individuals who have been instrumental in my pursuit of knowledge.

Foremost, I extend my deepest thanks to my professor, Huynh Van Nam, who has been an exceptional mentor and guide throughout an incredible eight-year journey encompassing my time as a research student, through my Master's, and culminating in my PhD. His wisdom, patience, and dedication have profoundly shaped my academic and personal growth. Professor Nam's unwavering support and insightful guidance have been a constant source of inspiration and strength. His ability to challenge and yet support me in equal measure has been instrumental in shaping not just my academic pursuits but also my approach to problem-solving and critical thinking in broader aspects of life.

Special thanks are due to Dr. Dinh Duy Tai, whose invaluable support and mentorship have been a cornerstone of my academic achievements. Dr. Tai's role transcended that of a mentor; he was instrumental in guiding me through the intricate process of academic submission and revision. His dedication went beyond teaching, as he generously took the time to personally revise and refine my journal paper. Without his meticulous guidance and hands-on assistance, completing my journal paper would have been a far more daunting task.

I am also deeply grateful to Dr. Nguyen Ba Hung for his significant contributions to my research and academic development. Dr. Ba Hung not only introduced me to the use of the JAIST supercomputer but also guided me into the field of Natural

Language Processing. His innovative ideas and insights were the catalysts that allowed me to merge my personal interests with academic research. His selfless assistance was particularly encouraging during the initial and often challenging phase of my PhD journey.

Above all, my deepest and most heartfelt gratitude is reserved for my wife, Fangyu Yu. Her sacrifice and unwavering support have been the cornerstone of my journey. Fangyu chose to leave a stable life in Singapore, resigning from her job, to join me in Japan, a decision that speaks volumes of her love and commitment. Our initial period here, especially during the COVID-19 quarantine, was filled with challenges. Yet, it was Fangyu's resilience and steadfast support that transformed these trials into a period of growth and togetherness. Her presence turned the hardships of quarantine into a journey of mutual support, strengthening our bond and reinforcing my resolve. Her sacrifices, her daily support in managing our lives and my academic endeavors, have been extraordinary. Fangyu has been more than a partner; she has been my pillar of strength, my source of inspiration, and the guiding light through the most demanding moments of my PhD journey. To her, I extend more than just my thanks; I extend a piece of my heart, filled with gratitude, love, and the promise of a future brightened by her presence.

Lastly, I extend my sincere gratitude to the Japan Advanced Institute of Science and Technology (JAIST). I am profoundly thankful for the comfortable and conducive environment that JAIST has provided, which has been instrumental in facilitating my study and research. The resources, support, and academic atmosphere at JAIST have not only enriched my learning experience but have also played a pivotal role in the successful completion of my dissertation. The institution's commitment to fostering a nurturing and innovative environment for its students is deeply appreciated and has significantly contributed to my academic growth and achievements.

To each and every one of you, I extend my profound thanks. This dissertation not only reflects my personal efforts but also stands as a testament to the collective support and wisdom imparted by each of you.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Introduction to Esports Industry

The rapid ascent of esports as a global entertainment giant is indisputable. The Newzoo 2019 Global Esports Market Report [1] indicated the industry's momentous achievement of surpassing $1 billion in revenue for the first time, propelled by an ever-expanding viewership that reached 453.8 million worldwide. This growth trajectory continued, despite global disruptions, as the 2020 [2] report suggested, with a remarkable increase in both audience size and revenue, even as the industry faced the shifting winds of global events and consumer behavior. By 2022, the esports industry was poised to achieve nearly $1.38 billion in revenue, with a global audience expected to reach 532 million, as noted in the 2022 report [3]. The same report highlighted the diversification of revenue streams beyond traditional sponsorships, with organizations venturing into blockchain, mobile esports, and direct-to-fan monetization strategies, such as digital merchandise and educational programs, thereby becoming resilient to potential shifts in the sponsorship landscape.

The foresight of esports entities in embracing technology to enhance fan experiences with innovations such as virtual reality spectator modes and augmented reality during the pandemic period has set a precedent for the integration of digital experiences with physical events. This integration promises to solidify esports' position at the vanguard of entertainment innovation. Moreover, the embracing of mobile esports and the exploration of blockchain technology for revenue generation speak to the industry's adaptability and forward-thinking approach. Esports organizations have also capitalized on the trend of lifestyle branding, with teams such as LOUD and 100 Thieves pioneering partnerships with fashion brands and expanding into educational and loyalty programs. These initiatives not only broaden revenue channels but also deepen fan engagement and the potential for brand partnerships.

Despite its remarkable expansion, a central challenge within the esports industry is its evolving market dynamics and revenue models, which present both opportunities and vulnerabilities in its economic structure. The market's traditional reliance on sponsorship, which accounted for 63% of the revenues in 2021, indicates potential vulnerability to economic fluctuations and changes in consumer behavior. This reliance is further compounded by the pandemic's impact, which has underscored the urgency for diversifying revenue streams beyond event-based earnings. In response, the industry has witnessed a shift toward digitalization with direct-to-fan business models, including digital merchandise, loyalty programs, and educational initiatives. Notably, the emergence of esports organizations in the public financial market, such as OverActive Media, Guild Esports, and Astralis, evidences the sector's maturation and the evolving interplay with traditional investment avenues. The current landscape presents a transformative phase for esports, characterized by its push into mobile gaming dominance in emerging markets as new revenue. These models have not only fortified revenue streams but also signified a transition towards esports organizations becoming multifaceted lifestyle brands, underscoring the sector's shift from a singular focus on gaming to broader brand positioning [3].

Over the last two decades, esports has had a global impact. It has evolved from a hobby, for players to a spectator sport, largely thanks to platforms like Twitch that enable live streaming. Nowadays tournaments fill arenas of traditional sports events and their prize pools often surpass those of conventional sports competitions. The world of esports is ever-changing with various gaming genres gaining popularity including Real Time Strategy (RTS) Multiplayer Online Battle Arena (MOBA) First Person Shooters (FPS) and Battle Royale games. Esports has transcended entertainment. Become a phenomenon that involves athletes, teams, coaches, commentators, and a passionate fan base. At present esports stands at the crossroads of technology, community engagement, and competitive spirit—a realm brimming with possibilities for the future. This transitional phase wherein esports is transitioning into a mainstream industry provides a context to explore the genres that have shaped its history while understanding key distinctions, between gaming and esports.

1.2 Challenges in Esports Environments

In the dynamic world of esports, novice players face a unique set of challenges that can impact their gaming experience. From navigating complex game mechanics to engaging with online communities, these hurdles are pivotal in shaping a newcomer's journey in competitive gaming. This section highlights challenges in esports as identified by existing research, focusing on the experiences of novice players in the competitive gaming landscape.

1. **High Learning Curve and Complexity**: Novice players entering the esports arena often confront a steep learning curve, particularly in games with intricate mechanics and strategic depth. This challenge is prominently observed in games like Dota2 or League of Legends, where understanding each character's unique abilities and the interplay between different game elements is essential. Cheung and Huang's [4] highlights the difficulties newcomers face in grasping complex game subtleties. Additionally, Bányai et al. [5] in their systematic literature review discuss the psychological aspects of esports, focusing on the cognitive and emotional challenges novice players encounter as they navigate through these complex gaming environments. This can be relevant in understanding the challenges faced by novice players in terms of game complexity and learning.

2. **Balancing Competition with Inclusivity**: The competitive nature of esports, while central to its appeal, often overshadows the need for inclusivity, especially for novice players. T.L. Taylor's [6] work delves into the gap between experienced and novice players, discussing how this disparity can affect newcomers' integration into the gaming community. Furthermore, Bányai et al. [5] also delve into aspects of esports that impact the inclusivity and accessibility of games, underscoring the importance of creating environments that cater to players of all skill levels.

3. **Toxic Teamwork and Community Dynamics**: The issue of toxicity within online gaming communities significantly affects novice players, who may be more vulnerable to negative behaviors such as harassment or exclusion. Kou and Gui's [7] study emphasize how toxic interactions can negatively impact player experience and retention, particularly in team-oriented games. This is further supported by Macey and Hamari [8], who explore the broader social implications of gaming and spectating in esports, highlighting the challenges of fostering positive and supportive online communities.

4. **Adapting to Game Updates**: Esports games are dynamic, with frequent updates and shifts that can profoundly affect gameplay and strategies. For novice players, keeping pace with these changes can be overwhelming. Seo and Jung's [9] research explores how rapid game changes require continuous adaptation, a significant hurdle for new players. Hamilton et al. [10] also address this issue, focusing on the role of community support in helping players navigate and adapt to these game updates, underscoring the importance of participatory and supportive gaming communities.

A predominant challenge in the esports domain is the insufficient structured understanding of players' experiences. While the difficulties faced by new players are acknowledged, the overwhelming volume and unstructured nature of player-generated content, including reviews and feedback, form a substantial barrier. This lack of structured analysis impedes the capacity of game developers, community managers, and stakeholders in formulating effective responses to these challenges. However, the innovative aspect-based sentiment analysis framework proposed in this dissertation successfully bridges this gap. By systematically analyzing vast quantities of player data, the research provides structured insights into players' experiences, directly addressing their specific needs and concerns. This groundbreaking approach significantly enhances the ability of industry professionals to make informed, targeted improvements, thereby elevating the overall esports gaming experience for new players."

1.3 Research Objectives

The primary objective of this dissertation is to address the aforementioned challenge by proposing an automated methodology for analyzing and interpreting the copious amount of game reviews generated by novice players. At the core of this novel approach lies a hybrid technique that strategically combines elements of topic modeling and sentiment analysis.

Topic modeling, a type of statistical model, is employed for discovering the abstract "topics" that occur in a collection of documents. On the other hand, sentiment analysis is a computational study of people's emotions, opinions, appraisals, and attitudes towards entities such as products, services, organizations, individuals, issues, events, and their attributes. By leveraging these two potent methodologies, our approach aspires to delve into the experiences, concerns, and sentiments of novice players, extracting crucial insights that often remain buried within unstructured text data. The primary aim of this research is to equip game operators with a detailed and nuanced understanding of their novice player base. By employing an automated methodology to analyze and interpret extensive game reviews, the study seeks to provide actionable insights. These insights are intended to directly inform and significantly enhance the gaming environment for novice players. While the immediate objective of this dissertation is to refine game development strategies for novice players, the broader, long-term ambition is to create a ripple effect that leads to a more balanced, inclusive, and engaging gaming experience for all. This research aims to lay the groundwork for promoting an equitable and sustainable esports ecosystem, envisioning a future where these improvements contribute significantly to the overall health and growth of the industry. To demonstrate the feasibility and efficacy of this proposed method, it will be applied to three representative PC esports games: Dota2, PUBG, CS:GO; and one mobile version: PUBG Mobile. The choice of these games was made considering their popularity, diverse player base, and distinct gaming mechanics, thereby enabling a comprehensive evaluation of our method across different types of games and gaming communities.

The significance of this research lies in its direct applicability to and potential impact on the esports industry, which is grappling with rapid growth and the need for strategic adaptation. This study's importance is multifaceted:

- For Academic Advancement: It pushes the envelope of existing knowledge in sentiment analysis by tailoring and testing its methodologies within the unique context of esports, thereby contributing new insights to the field of data analytics and computational linguistics.

- For Industry Practice: By providing a more nuanced understanding of player feedback, the research aids game developers, publishers, and event organizers in making informed decisions that can enhance player experiences, drive engagement, and potentially influence market trends.

- For Economic Strategy: The development of a framework that can interpret complex player sentiment data offers a strategic advantage in a market that is increasingly dependent on digital revenue streams and fan engagement for growth.

- For Innovation and Trend Analysis: The research facilitates the tracking of evolving player sentiments in real-time, allowing the industry to stay ahead of trends and innovate in response to the shifting demands of a diverse gaming audience.

By addressing these objectives, the research stands to bolster the esports industry's ability to maintain and grow its audience, diversify its revenue streams, and innovate within its offerings, ensuring its sustainable development in an ever-competitive entertainment landscape.

1.4 Research Contributions

1.4.1 Methodological Contributions

The dissertation introduces innovative theoretical frameworks and develops a novel sentiment analysis methodology that integrates advanced natural language processing techniques, specifically tailored to the distinctive demands of esports game reviews. These methodological contributions can be summarized as follows:

- The dissertation proposes a pioneering aspect-based sentiment analysis framework designed to extract deeper insights from extensive game reviews within the esports domain. This framework integrates enhanced topic modeling and sentiment analysis techniques to provide agile and informed responses to market demands and player feedback. A hybrid methodology is introduced, strategically combining Latent Dirichlet Allocation (LDA) with a TransfomerLite AutoEncoder for topic modeling and the BERT model for end-to-end aspect-based sentiment analysis. Experimental results demonstrate the methodology's effectiveness in extracting meaningful insights from unstructured text data.

- To illustrate the feasibility and efficacy of the proposed method, the dissertation applies the hybrid methodology to three representative PC esports games (Dota2, PUBG, CS:GO) and one mobile version (PUBG Mobile). These games represent the most popular and widely played titles in esports, providing comprehensive evaluations across diverse gaming communities and mechanics.

- While there is a wealth of literature on deep learning and sentiment analysis, the topics explored and the methodology proposed in this dissertation remain novel and distinctive. To the best of our knowledge, there are no prior works in the literature that officially address the same topics or design a methodology similar to what is presented in this dissertation.

1.4.2 Thorough Analytical Results and Findings

- The experiments were conducted on large-scale datasets collected from the Steam platform, including three representative esports games (Dota2, PUBG, CS:GO), and six million reviews for PUBG Mobile from Google Play. This extensive collection of reviews has resulted in a comprehensive dataset that captures a wide range of player perspectives. These datasets are made publicly available to the research community, contributing significantly to the available data resources in the existing literature and fostering increased exploration within the field.

- The experiments were conducted carefully to reveal the superiority of the proposed methodology. A comparative analysis was performed to verify the effectiveness of the proposed frameworks in solving the problems addressed in this dissertation. The following scenarios were investigated:

+ The quality of the proposed methodology leveraging LDA in combination with a Transformer architecture was compared with a variation applying only LDA. Results (in chapter 6) indicate that the proposed model offers a more nuanced understanding of the underlying topics within game reviews.

+ The proposed BERT-Transformer model's performance was compared with a variation using LSTM, confirming the superiority of the BERT-Transformer model (in chapter 6).

+ A comparison was made between the proposed model leveraging annotation fine-tuning and a variation without annotation fine-tuning. Results (in chapter 6) indicate that the model with annotation outperforms the model without annotation.

Overall, through a meticulous investigation, the dissertation has substantiated the logic and efficiency of the proposed methodology, showcasing its ability to accurately classify sentiments.

- The analysis results have been translated into easily understandable visualizations, shedding light on the diverse landscape of player sentiments. These visualizations facilitate a clear understanding of the range of opinions and experiences within esports communities, supporting game providers in enhancing their services and offering improved feedback for future development.

- The findings address numerous challenges in the esports landscape, identifying and tackling issues such as the high learning curve, inclusivity, toxic community dynamics, and adapting to frequent game updates. The dissertation illuminates the unique struggles of new players, providing valuable insights for game developers and community managers to enhance the overall gaming experience.

1.5 Dissertation Outline

This dissertation is divided into several chapters, each focusing on a unique aspect of the research. This approach ensures a comprehensive analysis and deeper understanding of the methodology proposed, particularly in its application to the esports context.

Chapter 1: This chapter introduces the burgeoning world of esports, detailing its rapid growth and unique challenges. It sets out the research objectives, provides a clear direction for the study, and concludes with an overview of the dissertation's structure.

Chapter 2: A comprehensive review of the existing literature is presented here, focusing on aspect-based sentiment analysis, topic modeling techniques, and the role of game reviews in the esports context.

Chapter 3: The chapter provides an in-depth exploration of the esports industry, highlighting the different genres and delving into the specifics of the selected games for this study, including PUBG, Dota2, CS:GO, and PUBGm.

Chapter 4: Here, the process of dataset preparation is elaborated, including the collection of data from Steam and Google Play, the methodology of text preprocessing, and the ethical considerations involved in the data handling process.

Chapter 5: This chapter outlines the proposed methodology for analysis, detailing the workflow of data collection, preprocessing, topic modeling, and sentiment analysis, with special emphasis on Latent Dirichlet Allocation (LDA) and End-to-End Aspect-based Sentiment Analysis by using BERT model.

Chapter 6: This chapter presents a comprehensive analysis of the datasets collected from both PC and mobile gaming platforms, with a focus on prominent titles such as PUBG, Dota2, and CS:GO for PC, and the mobile version of PUBG (PUBGm). The analysis delves into the nuances of topic modeling and sentiment analysis, drawing comparisons between the PC and mobile gaming experiences. It discusses the methodologies employed, interprets the sentiment distribution, and contrasts the findings across various player ratings, all while providing visualized data to support the insights. This integrated approach offers a holistic view of the gaming landscape, acknowledging the unique aspects of gaming platforms and their respective player communities.

Chapter 8: The final chapter brings together the key findings of the study, discusses the limitations encountered, and outlines potential avenues for future research in the field of esports.

Chapter 2

Literature Review

2.1 Aspect-based Sentiment Analysis

Aspect-based sentiment analysis (ABSA) emerges as a sophisticated analytical tool, pivotal in dissecting and quantifying public sentiment regarding distinct facets of products or services, drawing from the wealth of user-generated content on online discussion platforms. Within the dynamic world of esports, ABSA becomes particularly instrumental, offering granular insights into players' perceptions and emotional responses concerning diverse game elements, ranging from the visual appeal of graphics, the compelling nature of the storyline, to the quality of player interactions and community sociability. For example, players' feedback such as "the game graphics are stunning" underscores a positive sentiment, whereas references to prevalent cheating strike a negative chord.

Historically, the ABSA paradigm has been bifurcated into two critical tasks: aspect extraction and aspect sentiment classification. Aspect extraction delves into pinpointing the specific game features that resonate with players, employing both supervised learning techniques, which rely on pre-labeled data, and unsupervised methods that leverage algorithmic intelligence for pattern discovery. In a landmark contribution, Pontiki et al. [11] unveiled a diverse array of annotated datasets, spanning seven distinct domains and incorporating eight languages, thereby catalyzing multilingual ABSA research. Complementing this, Yauris et al. [12] innovated with a refined Double Propagation technique, adept at clustering aspect terms, thereby facilitating the generation of concise, aspect-oriented summaries. Adding a different dimension, Straat et al. [13, 14] ventured into the realm of consumer psychology, employing a meticulous feature extraction methodology anchored in word frequency analysis, coupled with manual polarity assessment, to gauge consumer sentiment in the gaming industry. Alturaief et al. [15]contributed to a review dataset AWARE: ABSA Warehouse of Apps Reviews, an ABSA dataset containing reviews of smartphone apps, and highlighted the use of supervised ABSA in the context of app reviews.

In a contemporary context, the ABSA field has witnessed transformative advancements. Baowaly et al. [16] pioneered a novel gradient-boosting framework, tailored for the nuanced analysis of Steam game reviews, distinguishing constructive criticism from unproductive negativity, and ingeniously predicting review scores through a sophisticated regression model. Parallelly, there has been a surge in deep learning applications [17, 18], harnessing the power of neural networks, although they necessitate substantial annotated corpora for effective model training.

Despite the strides in individual ABSA methodologies, they often fall short in capturing the multifaceted nature of player feedback, potentially skewing classification accuracy. This realization has spurred the evolution of a more integrative, robust approach known as End-to-End Aspect-based Sentiment Analysis (E2E-ABSA). This revolutionary model melds aspect extraction with sentiment classification into a cohesive analytical engine, delivering comprehensive, precise sentiment mapping. E2E-ABSA stands out in its mission to unveil not just the aspects players discuss but also the emotional nuances coloring these discussions. Ma et al. [19] were at the forefront, architecting an E2E-ABSA methodology that synergized the strengths of Long Short-Term Memory networks (LSTM), Convolutional Neural Networks (CNN), and Conditional Random Fields (CRF), setting a new benchmark in sentiment analytics. In a comparative study, Schmitt et al. [20] juxtaposed the efficacies of various neural network architectures, including LSTM and CNN, underscoring their potential in ABSA endeavors. Zhang et al. [21] propose a deep-level semi-self-help sentiment annotation system based on the bidirectional encoder representation from transformers (BERT) [22] weakly supervised classifier to explore consumer feelings. Urriza et al. [23] used Support Vector Machine (SVM) Classifiers to classify game reviews into Positive, Neutral, and Negative sentiments. Gunathilaka et al. [24] introduced a

CNN-based approach to analyze user reviews using Aspect-based Sentiment Analysis (ABSA).

In a significant development, Li et al. [25] ventured into optimizing the BERT framework specifically for E2E-ABSA, experimenting with alternative downstream layers to enhance the model's output precision. This exploration into custom-tailored neural network layers for ABSA signifies a critical step toward personalized sentiment analytics solutions. Furthermore, Agarwal et al. [26] showcased the versatility of the end-to-end model outside traditional reviews, extending its applicability to real-time social media sentiment tracking on platforms like Twitter. The beauty of E2E-ABSA lies in its streamlined efficiency; it eschews the need for labor-intensive feature engineering and intricate data preprocessing, marking it as a highly pragmatic, future-ready approach for real-world applications.

2.2 Topic Modeling

In the realm of digital data, the Latent Dirichlet Allocation (LDA) algorithm stands as a monumental innovation, particularly in the field of text analytics. Introduced by Blei et al. [27], LDA has revolutionized our ability to sift through extensive textual corpora, effectively illuminating the underlying topics that remain otherwise veiled in a sea of words. This unsupervised machine learning technique has found profound implications across various domains, demonstrating its versatility and efficacy in gleaning nuanced insights from raw, unstructured data.

One of the notable applications of LDA is in the analysis of online reviews, a domain replete with rich, qualitative user feedback. Huang et al. [28] leveraged LDA's prowess to dissect Yelp reviews, bringing to light diverse themes that customers frequently touch upon, such as the quality of service, cost-effectiveness, ambiance aesthetics, and health standards in restaurants. Their work underscored how establishments could harness these insights for qualitative improvements and targeted marketing strategies. The research conducted by Samy et al. [29] applied LDA and the class association rule Apriori algorithm for sentiment analysis of app reviews. Joung et al. [30] proposed an automated keyword filtering technique based on LDA to identify product attributes from online customer reviews. Guo et al. [31] proposed a co-branded potential LDA model for analyzing multiple clusters simultaneously to understand the strengths and weaknesses of several competing brands.

Extending LDA's reach into the financial sector, Feuerriegel et al. [32] embarked on an intriguing exploration of corporate communication. They applied LDA to corporate press releases, unraveling the thematic structures within and subsequently gauging their ripple effect on stock market performance. This pioneering work bridged the gap between textual content and market psychology, offering corporations a lens into the indirect financial impact of their public disclosures. Edison and Carcel [33] applied LDA to analyze the transcripts of the U.S. Federal Open Market Committee (FOMC), comprising 45,346 passages of text, over the period 2003-2012 to detect the evolution of the different topics discussed by FOMC members.

In the gaming arena, Santos et al. [34] utilized LDA to delve into the world of video game reviews. They unearthed striking differences in the thematic concerns of critiques penned by expert reviewers versus amateur commentators. This dichotomy highlighted the diverse lenses through which a game is evaluated, reflecting broader market segments and preferences. Wang and Goh [35] proposed a method to generate game experience dimensions by automatically processing online reviews of video games using an LDA algorithm. They identified seven meta-topics that included the main components of the gaming experience that players considered most important in their reviews: achievement, narrative, social interaction, social impact, visuals/value, accessories, and general experience. The results found that narrative and achievement were the main factors influencing player satisfaction with video games.

The consumer goods sector, too, has been ripe for LDA applications. Heng et al. [36] conducted a comprehensive analysis of Amazon food product reviews, identifying four predominant factors shaping consumer feedback: the efficiency of Amazon's service, the physical attributes of the product, taste profiles, and consumers' personal narratives or emotional expressions. This study demystified the drivers behind purchase decisions and brand loyalty, providing invaluable guidance for product development and marketing. Poushneh and Rajabi [37] used LDA techniques to categorize customer reviews and understand the relationship between reviews and their corresponding numerical ratings.

The travel and hospitality industry has also benefited immensely from LDA's analytical capabilities. Putri et al. [38] employed LDA to distill the essence of travel reviews, extracting thematic elements that significantly influence travelers' experiences and decision-making processes. In a similar vein, Tran et al. [39] analyzed a corpus of TripAdvisor reviews, revealing 11 dominant themes that stood out in travelers' discussions, encompassing culinary experiences, accommodation amenities, pricing, and staff behavior, among others. Sutherland et al. [40] utilized a large amount of unstructured text data from 104,161 online reviews from Korean lodging patrons and used an inductive approach to identify topics of interest that guests consider important. Afaq et al. [41] adopted two powerful techniques, topic modeling, and sentiment analysis to analyze the concerns and emotions of hotel customers towards the South Asian global hotel chain.

These diverse applications of LDA across sectors underscore its transformative potential in harnessing the subjective human experience, converting qualitative narratives into quantifiable data. This data, rich with insights, empowers businesses and researchers to comprehend and cater to the intricate tapestry of human needs and preferences better.

2.3 Game Reviews

In the digital era, game reviews have emerged as a critical facet of user-generated content, offering a wealth of insights into player preferences, experiences, and expectations. Unlike more restrictive platforms like Twitter, gaming forums and review platforms typically impose minimal limitations on content length or format. This liberal approach, while encouraging detailed feedback, also introduces challenges such as inconsistent data quality, the presence of non-standard text, and susceptibility to spam and irrelevant content.

The open format of game review platforms fosters a rich tapestry of linguistic expressions, encompassing everything from everyday colloquialisms to specialized gaming terminologies. This diversity presents a unique and fascinating challenge, especially in the esports context. Our research embraces this complexity head-on, utilizing advanced computational techniques that go beyond mere word replacement or review trimming. Instead, our approach preserves the authenticity and cultural richness of player feedback, while skillfully deciphering the varied sentiments and nuanced content embedded within. This method represents a significant stride in understanding and valuing the full spectrum of player expressions in their original form, reflecting the true diversity of the gaming community.

Over the years, scholarly exploration in the realm of game reviews has been extensive. Gifford [42] embarked on a comparative study of reviews across different entertainment mediums, highlighting the unique evaluative criteria employed in video game reviews compared to film critiques. This work shed light on the distinct cultural and contextual factors influencing consumer feedback in the gaming industry. Busurkina et al. [43] discovered how people evaluate their gaming experience based on reviews on the Steam platform and extracted its main dimensions. Pielka et al. [44] extracted potentially influential topics by analyzing online game reviews posted on a German gaming network and finding that these reviews provided detailed descriptions of gameplay and technical and playful issues.

In a pivotal study, Lin et al. [45] underscored the disparity between game reviews and mobile app feedback. They revealed that the gaming community often provides constructive criticism, with both accolades and complaints offering valuable operational insights to game developers and marketers. This nuanced understanding of reviews goes beyond star ratings, delving into the specific attributes that engender player satisfaction or discontent.

Zagal et al. [46] ventured into the psycholinguistic aspect of game reviews, illustrating how the choice of sentiment-laden words correlates with overall game ratings. Their research provided a psychological perspective on how emotional expression in reviews reflects player attitudes and experiences.

Further enriching the field, Bond et al. [47] conducted an in-depth analysis of game reviews to distill the characteristics that define a “good game”, offering a guideline for developers aiming for commercial and critical success. Concurrently, Livingston et al. [48] established a direct correlation between the tenor of game reviews, the ratings they receive, and their subsequent market performance, emphasizing the influence of public opinion on sales and profitability. Li et al. [49] analyzed the playability of a video game based on player reviews. The results show that a data-driven approach provides an effective solution for understanding the overall playability of a game as well as specific strengths and weaknesses at the game system level.

There are several works utilizing Long Short-Term Memory for game reviews. Ghosh et al. [50] presented an approach to addressing the lack of research on real-time detection of emotional states in hardcore gamers during Android gaming applications. The proposed design focuses on utilizing electroencephalographic (EEG) signals for this purpose, incorporating a four-step process of data collection, pre-processing, feature extraction, and classification. The proposed framework uses an attention-based Bi-directional Long Short-Term Memory (Bi-LSTM) network to classify emotional states into five categories: happiness, sadness, surprise, anger, and neutral. Sivakumar et al. [51] proposed a framework for sentiment analysis in Amazon video game reviews, utilizing Long Short-Term Memory (LSTM) with fuzzy logic. This approach enables the classification of consumer review sentences into four distinct labels: highly negative, negative, positive, and highly positive. By leveraging the strengths of LSTM, the proposed model achieves an accuracy of 83.82% when applied to Amazon video game reviews.

Hermans [52] tried to answer the question "how can we create legitimate positive reviews for video games using unsupervised processing methods and AI". Several techniques are employed to address the research question. Initially, the data undergoes filtering to include only positive reviews, subsequently undergoing a division into sentences for further review filtering based on sentence count. Further refinement involves selecting only English reviews, followed by a segmentation of reviews into lists of words for model utilization. To establish a baseline, an n-grams model is implemented for comparison with the neural network. The neural network, a sequential Keras model, comprises an embedding layer, an LSTM layer, and a dense layer. Both models undergo scrutiny to identify instances of copying from the training set through cosine similarity checks. The comparison between the two models encompasses run-time and review quality assessments to determine the superior model. Jiaxin Song [53] designed an aspect-based sentiment analysis approach for mobile game reviews using deep learning. The method leverages extensive mobile game review data to assess users' emotional tendencies across different attributes of the game with fine-grained precision. The sentiment analysis method consists of three models: a baseline model incorporating Bi-LSTM, a Fully Convolutional Network (FCN), and a Conditional Random Field (CRF) for sentiment collocation extraction, matching, and classification.

2.4 Summary

Recent advancements in aspect-based sentiment analysis (ABSA) and topic modeling, particularly in the context of digital media analytics, have opened new avenues for understanding user-generated content. Studies in ABSA have evolved from basic sentiment detection to more nuanced interpretations, yet they often fall short of fully capturing the complexities of player feedback in the gaming industry. This is further compounded in the realm of topic modeling, where the Latent Dirichlet Allocation (LDA) algorithm has shown its potential across various domains. However, the unique linguistic nuances and evolving jargon of the gaming community present a significant challenge, underscoring the need for more tailored and sophisticated analytical approaches.

In parallel, extensive research into game reviews has revealed their critical role in offering insights into player preferences and experiences. These reviews, however, are characterized by diverse and unrestricted linguistic expressions, making traditional text analysis techniques less effective. The current literature emphasizes the distinctive nature of game reviews compared to other consumer feedback forms, highlighting their untapped potential in guiding game development and market strategies. Our research aims to bridge these gaps by integrating ABSA and LDA methodologies, creating a tailored solution that can effectively process and analyze the complex and dynamic nature of game reviews. This innovative approach not only promises to enhance the accuracy and depth of esports player feedback analysis but also contributes significantly to the broader understanding of digital sentiment analysis and text analytics.

Despite the wealth of research, a noticeable gap persists in the nuanced study of esports game reviews. Most analyses treat esports similarly to traditional gaming, overlooking the unique community dynamics, competitive nature, and distinct player expectations inherent to esports. Recognizing this oversight, our study proposes a dedicated exploration of esports game reviews using an advanced End-to-End Aspect-based Sentiment Analysis (E2E-ABSA) model, grounded in the BERT architecture [25]. By training this model on a curated dataset from the vibrant Steam community, we aim to unveil the intricate player opinions hidden within esports reviews. The insights garnered promise to equip esports service providers with the nuanced player feedback necessary to refine their offerings and elevate the player experience.

Chapter 3

Background

3.1 Introduction of Esports

Electronic sports, or esports, have undergone a remarkable transformation over recent decades. Initially a niche subculture characterized by small, local network gatherings, esports has grown into a global industry worth billions, drawing in millions of viewers and revolutionizing both entertainment and competitive sports. This evolution has elevated esports beyond mere gaming, turning it into a significant cultural movement. It has led to the emergence of professional players, teams, coaches, and commentators, creating an ecosystem similar to traditional sports. Universities are now offering esports programs and scholarships, and cities worldwide are investing in esports arenas. This growth of esports is not only indicative of technological progress but also mirrors a shift in societal perspectives on gaming. What was once seen as a solitary, somewhat isolating pastime has become a means of fostering community, social engagement, and even a viable career for some.

The embryonic phase of esports can be traced back to 1972, where Stanford University held the first recorded video game competition for the game Spacewar. These early years laid the foundation of competitive gaming as an appealing, engaging activity. The late 1980s and early 1990s brought about the era of arcade gaming, catapulting video games to broader societal awareness. Games like Street Fighter II emerged, instigating a new wave of enthusiasm for competitive gaming through local tournaments.

The late 1990s witnessed a shift, in the video game industry with the advent of the internet, which marked an era and a crucial milestone in the history of esports. During this time players from over the world engaged in time online competitions breaking through geographical barriers and driving the growth of competitive gaming. Games such as Quake, Starcraft and Counter Strike played a role, in shaping the esports scene during that period. As we entered the millennium video game technology continued to evolve giving rise to esports. Enhanced graphics, more complex gameplay, and the development of engaging virtual environments all played significant roles in advancing esports. It was during this time that major esports tournaments, such as the Electronic Sports World Cup in 2003 and Major League Gaming in 2002, were established, offering gamers a global platform to demonstrate their prowess.

3.1.1 Genres of Esports

Esports includes a range of video game genres each bringing its distinct element to the world of competitive gaming. These genres differ in terms of gameplay mechanics, objectives and styles. Frequently serve as the foundation, for individual esports tournaments.

Real-Time Strategy (RTS) e.g., *StarCraft II, Age of Empires IV*

RTS games challenge players with real-time decision-making and resource management in a dynamic environment. Players must build and manage resources while engaging in strategic combat. Titles like StarCraft II and Age of Empires IV are notable for their complex strategies and depth, making them staples in competitive gaming.

Multiplayer Online Battle Arena (MOBA) e.g., *Dota2, League of Legends*

In MOBAs, players take command of a character as part of a team with the objective of obliterating the enemy teams base. Games such as Dota2 and League of Legends demand teamwork. The skillful utilization of distinct hero abilities, which has earned them a significant following in the world of esports due, to their exhilarating and strategic gameplay.

First-Person Shooters (FPS) e.g., *Counter-Strike: Global Offensive, Overwatch*

FPS games, such as Counter-Strike; Global Offensive, and Overwatch provide an exhilarating and challenging gaming experience. These games require aiming, strategic coordination, among teammates and fast reflexes. They are highly praised in the esports community for their strategic gameplay.

Battle Royale e.g., *PlayerUnknown's Battlegrounds (PUBG), Fortnite*

The Battle Royale genre, with games like PUBG and Fortnite, features a survival-based, last-player-standing format. Players must scavenge for resources and engage in combat within a shrinking play area, making these games thrilling and unpredictable.

Fighting Games e.g., *Street Fighter, Tekken*

Fighting games revolve around face, to face combat within a controlled setting. Players choose characters with sets of moves. Participate in battles that rely on skill. Titles such as Street Fighter and Tekken are renowned for their one, on one showdowns demanding players to master timing and strategic thinking.

Sports Games e.g., *FIFA, NBA 2K, Madden NFL*

Sports games such as FIFA and NBA 2K provide an experience of real-world sports enabling players to strategize and control teams. These games are highly regarded for their authenticity. Frequently include licensed teams and players, from the world, which makes them appealing to both sports enthusiasts and gamers alike.

Each genre offers unique strategic depths and competitive elements, contributing to the rich diversity of esports. They cater to different player preferences, skills, and strategies, thus enhancing the inclusivity and reach of esports globally.

3.1.2 Difference between Games and Esports

Video games have become a part of our culture for years starting from the early days of arcades and home consoles in the 1970s and 1980s [54]. Over time video games have evolved into an expansive medium ranging from mobile games to complex massively multiplayer online games (MMOs). The main purpose of video games is to engage and entertain players offering experiences. Gaming encompasses everything from single-player narrative adventures to multiplayer environments where players can connect and compete in a realm. Within this broad spectrum of gaming, esports represents a specialized niche. Esports are defined by their focus on competitive and structured gameplay, setting them apart within the larger gaming industry. Much like traditional sports, esports involve competition between players and teams, strict rule sets, tournaments, leagues, and spectators. Esports have seen a surge in popularity over the past decade, driven by advancements in internet connectivity, live streaming technology, and the growing legitimacy of professional gaming [55].

The primary difference between video games and esports lies in the level of competition and organization. While competition can exist in video games, not all competitive games qualify as esports. For a game to be considered an esport, it must support fair and balanced competition among its players. In addition, there typically needs to be an organized structure for competitive play, including tournaments, leagues, or ladder systems. Furthermore, not all video games are designed with esports in mind. Many games are created for casual play, with an emphasis on narrative, exploration, or social interaction rather than competitive gaming. Other games, while competitive in nature, may not have the necessary balance or structure to support an esports scene [56]. On the other hand, esports titles are often designed or adapted to promote high-level competition. They generally have a high skill ceiling, meaning that a player can spend a significant amount of time improving and still find new strategies or techniques to learn. Esports also have spectator-friendly elements, as a significant part of esports' appeal comes from the ability to watch highly skilled players compete.

Essentially while all video games fall under the category of esports, not all video games can be classified as esports. This distinction relies on factors such, as the level of competitiveness the presence of competitive structures and the aspect of spectatorship[57]. It's important to grasp these differences but important to acknowledge the shared aspects between video games and esports. Both can cultivate communities offer players a sense of achievement and serve as a form of entertainment. Moreover recognizing these disparities can provide insights, into the experiences, challenges, and expectations that players may have based on whether they are engaging in casual gaming or competitive esports.

3.2 Introduction to selected esports games

3.2.1 Defense of the Ancients 2 (Dota2)

Defense of the Ancients 2 (Dota2) exemplifies the depth and complexity found in the MOBA genre. Created and published by Valve Corporation, it has been a forefront in the world of esports since its release in 2013. This game features a robust and intricate gameplay experience, where strategic planning, real-time decision making, and team coordination converge to form a highly competitive environment, as demonstrated in Fig. 3.1.

Figure 3.1: Dota2 gameplay screenshot.

At its core, Dota2 revolves around achieving objectives engaging in team fights and managing resources. Two teams consisting of five players each known as the Radiant and the Dire battle it out to destroy their opponents "Ancient", a structure situated within their base. The map is symmetrically divided by a river, with three "lanes" serving as paths to the enemy base, each guarded by defensive turrets and spawning waves of units called "creeps" that assist the players.

Dota2's strategic diversity stems from its collection of over 100 "heroes", each possessing abilities and fulfilling specific roles within a team. These heroes are grouped into categories like "carry", "support", "offlane", "mid" and "jungler", with each category having responsibilities. Mastering a hero not only requires understanding their skills but also comprehending how they synergize with teammates and counter adversaries. One of the defining features of Dota2's gameplay is the concept of "fog of war". This mechanic restricts a teams vision to areas where they have units or structures adding a dimension to the game. Vision control becomes a tactical aspect of the game, as teams set up "wards" for sight or remove the enemy's wards to gain an advantage. Dota2's complexity is further amplified by its items system. Players earn gold by defeating creeps, heroes, and structures, which they can use to purchase items that enhance their heroes' abilities. The choice of items can significantly affect the outcome of a match, making each game as much a battle of wits as it is of reflexes.

Valve's commitment to the esports community surrounding Dota2 shines through its flagship tournament, *The International*. This tournament not only captures attention with its groundbreaking prize pools. Also serves as a platform to showcase the pinnacle of Dota2 gameplay. It unveils the level of maneuvers, team compositions and individual skills. Ultimately the enduring popularity of Dota2 in the esports realm can be attributed to its gameplay, diverse and balanced hero selection and Valve's dedication to fostering a competitive environment. It stands as an exemplar of the MOBA genre consistently evolving through updates and new hero introductions that ensure a captivating experience for both players and spectators.

3.2.2 Counter-Strike: Global Offensive (CS:GO)

Counter-Strike: Global Offensive (CS:GO), a cornerstone in the realm of competitive first-person shooters, has been a defining force in esports since its release in 2012. Developed by Valve Corporation and Hidden Path Entertainment, CS:GO has enjoyed a prestigious status in the esports world, thanks to its strategic depth and skill-based gameplay. This commitment to high-level competitive play is evident in Fig. 3.2, which captures the intense moment of a CS:GO match.

Since September 2023 CS:GO has gone through a significant update, with the launch of CS2. This update brings improved graphics and performance by transitioning to the Source 2 engine. However, it's worth mentioning that despite these upgrades the fundamental gameplay of CS:GO remains unchanged. It still maintains its depth. Relies on skill based mechanics. Therefore it is important to note that any analysis or discussions in this paper will continue refer to the game as CS:GO,

Figure 3.2: CS:GO gameplay screenshot.

considering the time when the research was conducted.

The gameplay of CS:GO involves two teams, the Terrorists and the Counter-Terrorists, facing off in various objective-based modes, the most prominent being bomb defusal and hostage rescue. The game is celebrated for its precise shooting mechanics, strategic map designs, and the crucial role of team tactics. The in-game economy system, where players earn and spend money on weapons and equipment based on their performance, adds an additional layer of strategy to each round.

CS:GO's arsenal of weapons, characterized by distinct recoil patterns and usage techniques, contributes to its high skill ceiling. Mastery of these weapons, along with a deep understanding of the maps and sound tactical planning, is essential for success in the competitive arena. The game's map design is a key element in its strategic depth. Each map offers a unique layout requiring specific tactics and strategies, making map knowledge an invaluable component of high-level play. The balanced design ensures that both Terrorists and Counter-Terrorists have equal opportunities to execute their strategies.

The esports ecosystem of CS:GO is robust, featuring a range of prestigious tournaments, including the celebrated Majors. These events not only showcase the highest level of skill and competition but also attract a large global audience, solidifying CS:GO's status as a popular spectator sport. In summary, CS:GO remains a standard setter in the world of first person shooters. Its blend of gameplay, skill based mechanics and dynamic map strategies continues to captivate players and fans making it an iconic title, in the esports industry.

3.2.3 PlayerUnknown's Battlegrounds (PUBG)

PlayerUnknown's Battlegrounds (PUBG) emerged as a trailblazer in the esports industry by popularizing the "Battle Royale" genre. It offers a unique combination of survival tactics, strategic planning, and skillful play. The game's premise, which pits individuals and teams against each other in a shrinking play zone, is simple in concept but complex in its execution, resulting in a rich tapestry of gameplay nuances that are depicted in Fig. 3.3. PUBG's game design intricately balances elements of risk and reward, strategy and randomness, engagement and retreat. Players begin their journey airdropped onto an island, equipped with nothing, and must scavenge for weapons, vehicles, and supplies. They face a dual threat: other players and the environment itself, as the playable area contracts due to a toxic gas, ingeniously fostering encounters and battles as the game progresses.

Figure 3.3: PUBG gameplay screenshot.

The landscapes of PUBG are meticulously crafted, featuring a variety of terrains such as dense forests, open fields, and urban areas, which influence the tactical decisions players must make. Weather conditions and the time of day dynamically change in each match, adding layers of unpredictability and requiring players to adapt their strategies continuously. PUBG's arsenal is extensive, offering a wide range of firearms, melee weapons, and throwables, each with realistic ballistic properties. The game's attention to detail extends to the level of weapon customization, where players can equip attachments to suit their combat style, further enhancing the strategic gameplay that has become a signature of the game.

The social element of PUBG also contributes significantly to its success in esports. By offering options to play with a partner or in a team the game caters to social and competitive preferences. Effective communication and coordination, during team gameplay are crucial. Can ultimately determine the outcome during final moments of a match.

In summary, PUBG has established a benchmark, for the qualities that a Battle Royale game can provide. It offers not entertainment but a competitive experience of great value. The games authentic survival mechanics along, with its complexity and the intensity of high stakes matches guarantee that each round becomes a legend that captivates both players and onlookers.

3.2.4 PUBG Mobile (PUBGm)

PUBG Mobile (PUBGm) represents a significant milestone in the evolution of mobile gaming into a legitimate platform for esports. As the mobile adaptation of the acclaimed PC game: PUBG, PUBGm brings the intense and strategic gameplay of its predecessor to a wider, more accessible audience. This transition to mobile gaming has been captured in Fig. 3.4, showcasing the game's interface and mechanics adapted for touchscreen devices. Developed by Tencent Games in collaboration with PUBG Corporation, PUBGm was released on the Google Play Platform and quickly rose to prominence due to its faithful recreation of the PUBG experience.

The game features a similar battle royale format where 100 players parachute onto an island and scavenge for weapons and equipment to battle it out until only one player or team remains. Despite the constraints of mobile devices' performance, PUBGm manages to retain the core elements that made the PC version successful. It showcases impressive graphics for a mobile title, a wide variety of weapons, vehicles, and equipment, and a large, open-world map that shrinks over time, forcing players into increasingly close encounters.

Figure 3.4: PUBGm gameplay screenshot.

One of the critical aspects of PUBGm's success in esports is its accessibility. The game can be played on most smartphones, dramatically expanding the player base and fostering a diverse gaming community. This accessibility has made the game particularly popular in regions where mobile gaming is more prevalent than PC or console gaming. PUBGm's control scheme is meticulously designed to translate the complex actions of a battle royale game onto a touchscreen interface. This design includes customizable controls and aim-assist features that help mitigate the challenges of playing a shooter game on a mobile device, making the game approachable for both casual and competitive players.

The esports scene for PUBGm has seen substantial growth, with numerous tournaments being held worldwide. These events range from local and regional competitions to international championships, featuring large prize pools and attracting significant viewership, highlighting the game's global appeal.

In summary, PUBGm has successfully translated the battle royale genre to a mobile platform, creating an esports phenomenon that extends beyond traditional gaming setups. Its combination of accessibility, comprehensive gameplay, and competitive integrity makes it a standout title in the world of mobile esports.

3.3 Justification of selected esports games

In the multifaceted world of esports, our selection of Dota2, CS:GO, and PUBG, including its mobile version PUBGm, for research purposes is both deliberate and strategic. This choice encapsulates the varied dimensions of esports, from individual skill to team coordination, across different platforms and genres.

Transitioning to the complex world of MOBAs, Dota2 offers a stark contrast with its emphasis on team-based strategy and cooperative gameplay. Valve's flagship game has set a high bar in the esports industry with *The International*, an event that epitomizes the grandeur and global appeal of esports through its record-breaking prize pools and extensive international viewership. Dota2's ever-evolving gameplay, with a vast array of heroes and constant updates, presents an ideal case for studying the dynamics of team play, strategic evolution, and the impact of game updates on competitive ecosystems.

CS:GO, a paradigm of first-person shooters, brings to the forefront tactical gameplay, precision, and team synergy. The game's enduring presence in esports is marked by prestigious events like the *CS:GO Major Championships*, which draw elite teams and vast audiences worldwide. CS:GO's sustained popularity provides a lens through which to examine the evolution of esports communities, the sustained appeal of tactical shooters, and the elements that contribute to the longevity and vibrancy of esports titles.

PUBG and PUBGm together encapsulate the adaptability and expansiveness of esports. PUBG, pioneering the Battle Royale genre, introduced a new dimension of survival and strategy to competitive gaming, evidenced by tournaments like the *PUBG Global Championship*. The porting of PUBG to the mobile platform as PUBGm further broadened the genre's reach, as exemplified by the *PUBG Mobile Global Championship*. This porting is not merely a change in platform but also reflects the growing trend of mobile gaming in esports. The extensive user review data from Google Play for PUBGm enhances our ability to conduct a nuanced comparison between PC and mobile gaming experiences, offering insights into platform-specific dynamics, player engagement, and the evolving landscape of esports.

In essence, the selection of these games for our research offers a comprehensive view of the esports domain. From the strategic depth of Dota2's team play, the tactical precision in CS:GO, to the innovative and adaptive gameplay in PUBG and its mobile iteration, this diverse array of games provides rich datasets and varied contexts, making them ideal subjects for an in-depth exploration of the factors driving success, engagement, and evolution in the esports industry. This selection enables an in-depth exploration of the factors driving success, engagement, and the continual evolution within the vibrant and ever-expanding world of esports.

Chapter 4

Dataset Preparation

4.1 Introduction

The foundation of any data-driven research lies in the quality and integrity of its dataset. Dataset preparation, especially for text-based feedback like ours, involves handling inconsistencies, making decisions about granularity, and understanding its structure. This phase not only refines the data but also uncovers its patterns and potential biases, which can guide later stages of analysis. This chapter delves into the steps taken to prepare the dataset for this research, from initial data collection to the preprocessing techniques applied.

4.2 Dataset Collection from Steam

4.2.1 Steam Platform

Steam, launched by Valve Corporation in September 2003, rapidly ascended to prominence as one of the world's preeminent digital distribution platforms for PC gaming. Over the years, it has become a hub for gaming enthusiasts, developers, and researchers alike. The platform's vast catalog of games, ranging from indie titles to blockbuster hits, garners millions of reviews from its expansive user base. For this study, the decision to harness Steam as the exclusive source for data collection was informed by several compelling factors:

1. Volume of Reviews: With its vast library of games and an active user base, Steam is

a goldmine of player reviews. These reviews range from succinct praise or criticism to detailed evaluations of gameplay mechanics, graphics, audio, and other integral game elements.

2. Authenticity of Feedback: Steam users are known for their candidness. They not only rate games but also provide genuine feedback based on their personal experiences. These first-hand accounts provide rich insights, making them invaluable for academic research.

3. Diversity in Player Profiles: Steam's global reach ensures a heterogeneous mix of player feedback. From casual gamers to esports professionals, the reviews encapsulate varied gaming experiences, further enriching the dataset.

4. Ease of Accessibility: Steam provides structured access to game reviews via its Steamworks Web API, facilitating automated and systematic extraction.

5. Relevance to Esports: Given the rise of esports titles on Steam, such as Dota2 and CS:GO, the platform's reviews provide firsthand player experiences related to competitive gaming and its various facets.

Within our exploration of Steam as a source platform, we chose "PUBG: BATTLEGROUNDS" as a representative example to illuminate the structure and nuances of the platform's review system, as depicted in Figure 4.1. The screenshot underscores several intrinsic features of Steam's review mechanism. It highlights the platform's ability to filter reviews based on various criteria such as overall helpfulness and language—evidenced by options like "Most Helpful (All Time)" and the choice of "English". Each review is prominently labeled as either recommended or not, offering immediate insight into the player's sentiment. Complementing this is the inclusion of playtime and date, contextualizing the reviewer's depth of experience and the relevance of their feedback. The content of the reviews, ranging from specific concerns about privacy policies to broader gameplay experiences, showcases the breadth of topics players discuss. Furthermore, the interaction metrics, such as "helpful" and "funny" tags, emphasize the platform's community-driven nature. Overall, this review page epitomizes the richness and structure of data obtainable from Steam, providing us with a comprehensive insight into player sentiments and feedback.

Figure 4.1: Steam Review Snapshot for PUBG: BATTLEGROUNDS.

4.2.2 Data Collection

Utilizing the steamreviews Python Library

The steamreviews[1] module offers a rich interface for accessing the Steam reviews in a Python environment. Developed as an open-source utility, it simplifies the process

[1] https://pypi.org/project/steamreviews/

of fetching large quantities of data, especially for those not familiar with Steam's intricate API endpoints, such as the Steam API's inherent rate limit, which is approximately 10 reviews per second.

When considering data collection on a larger scale, where multiple games or software reviews might be of interest, the capability of the module to process a batch of appIDs is invaluable. By providing a list of unique appIDs—identifiers for every product on the Steam platform—researchers can obtain reviews for multiple games with a single command. This batch processing method is not only efficient but also reduces the chance of hitting API rate limits by optimizing the number of requests made. In the realm of customizable data collection, the library stands out. It allows users to specify particular parameters for review extraction, such as language or sentiment. For instance, a researcher interested in understanding the sentiment of non-English reviews can set the language parameter accordingly. Similarly, filtering reviews based on sentiment (positive or negative) can aid in sentiment analysis studies, shedding light on factors driving positive reception or common issues leading to negative feedback.

However, based on the outcomes of our experiments, we discovered that even when attempting to gather all reviews for a particular game, it's highly probable that we might miss some. Consequently, we suggest that when amassing reviews across multiple games, it's more effective to gather them individually rather than in bulk. Furthermore, we theorize that the Steam backend's limitations necessitate avoiding high-traffic periods, such as weekends and holidays, during the review collection process.

Direct Engagement with Steam's Review API

The official Steam interface is comprehensive, offering a dedicated endpoint for retrieving user reviews:

`GET store.steampowered.com/appreviews/<appid>?json=1`

This API is highly configurable, allowing reviews to be sorted based on diverse metrics, specified by language, filtered within a particular temporal range, paginated, categorized by sentiment, distinguished by purchase origin, limited in number per call, and even filtered for off-topic content. The output of this API is a detailed JSON structure. It encapsulates not only individual review details—like recommendation ID, SteamID, game count, review content, and timestamps—but also aggregate data, such as overall review counts and scores.

In this experiment, we configured the parameters as follows: The `language` was set to `en`, and `num_per_page` was adjusted to its maximum value of 100. This configuration was chosen to expedite the collection of all English reviews. It was gratifying to note that Steam has officially incorporated the `filter_offtopic_activity` option, which serves to exclude off-topic reviews, commonly referred to as `Review Bombs`. This invaluable feature was introduced subsequent to the data collection phase of our experiment. Its inclusion promises to greatly enhance the quality of datasets in our forthcoming studies.

Our data collection strategy revolved around three representative esports games: Dota2 (a classic example of MOBA), PUBG (an icon in the BR category), and CS:GO (a stalwart in the FPS category). The choice of these games was not arbitrary. Drawing on insights from industry survey reports, specifically Newzoo, and trends from significant esports events, these games stood out for their prominence in the esports scene and their substantial player and reviewer base. Utilizing Steam's API granted us the ability to meticulously crawl and capture game reviews. This methodology allowed us to collate a fresh dataset of English reviews, specific to the three selected esports games, extending up until December 2021.

In our in-depth study, while the Steam Review API offers a plethora of features for each game review, we deliberately chose to concentrate on a subset of these features. This strategic decision ensured that our analysis was streamlined, targeted, and relevant to the research questions at hand. The following are the primary features we considered pivotal for our research:

- **recommendationid**: Serving as the cornerstone of each review, the unique recommendation ID ensures that each piece of feedback is distinctive and traceable. This identifier is crucial to avoid duplicates and provide authenticity to each review.

- **playtime_at_review**: An essential metric that offers invaluable context, the playtime when the review was written provides insight into the depth of the user's experience with the game. This is pivotal in discerning whether feedback comes from a seasoned player or someone relatively new to the game.

- **review**: As the crux of our sentiment analysis, the text of the written review provides raw, unfiltered insights from players. This textual data is where themes emerge, patterns are recognized, and genuine sentiments are unveiled.

- **timestamp_updated**: Capturing the dynamism of user feedback, the date the review was last updated (expressed as a unix timestamp) reflects the evolving nature of player sentiment. It can also indicate patches, updates, or events that might have influenced a change in opinion.

- **voted_up**: A straightforward yet impactful metric, this binary feature quickly segregates reviews into positive and negative sentiments. If "true", it indicates a positive recommendation, streamlining the initial stages of sentiment analysis.

- **votes_up**: Beyond the content of the review, how it's perceived by the community adds another layer to its significance. The number of users that found a review helpful is a testament to its resonance within the player base and can point to prevailing community sentiments.

- **votes_funny**: Esports and gaming, while competitive, have a lighter side too. Capturing the number of users that found a review humorous gives a nuanced understanding of the community's temperament and what they find entertaining or ironic in game experiences.

4.3 Dataset Collection from GooglePlay

4.3.1 Google Play Platform

Google Play, which was introudced by Google in October 2008, swiftly became a dominant force in the digital distribution market for mobile applications. As a primary source for Android apps, games, and various digital content, it has attracted a vast and diverse user base. For researchers and developers, Google Play offers an opportunity for data collection, particularly in the realm of app reviews and ratings. The platform's extensive app repository and the wealth of user-generated content make it an ideal candidate for data-driven studies. The rationale behind selecting Google Play for data collection encompasses several key aspects:

1. Volume of Reviews: GooglePlay hosts millions of apps, each accompanied by user reviews and ratings. These reviews range from brief user impressions to detailed feedback on app functionality, usability, and user experience.

2. Authenticity and Diversity of Feedback: The reviews on Google Play are generated by a global user base, offering a wide spectrum of perspectives. This diversity is crucial for obtaining a holistic view of user sentiment and app performance.

3. Accessibility of Data: Google Play provides structured access to app reviews through its Google Play Developer API, enabling streamlined and efficient data extraction.

4. Relevance to Mobile App Market Trends: As a leading app store, Google Play's data is reflective of current trends and user preferences in the mobile app market, making it a valuable resource for market analysis and research.

In the competitive sphere of mobile gaming, the significance of player input is magnified, as it directly influences game development and fosters community relations. This dynamic is exemplified by the "PUBG Mobile" review page, as depicted in Figure 4.2, where a composite rating of 4.3 out of 5 stars, derived from millions of user ratings, underscores the game's widespread appeal and the active engagement of its players. The straightforward display of this rating, alongside the spread of individual star ratings, provides an instant read on the game's positive reception. Reviews on the page, detailed with user names, submission dates, and star ratings, offer a transparent lens into the players' experiences. Noteworthy contributions from users, such as Travis Bryson and Joseph Cross, range from critiques on technical glitches to praise for the game's equitable progression system, painting a vivid picture of its merits and shortcomings. The developer's responsiveness to such feedback is made evident through direct interactions with users, demonstrating a commitment to customer service that enhances the community's trust. Additionally, the platform's interactive element lets players assess the helpfulness of reviews, a community-driven

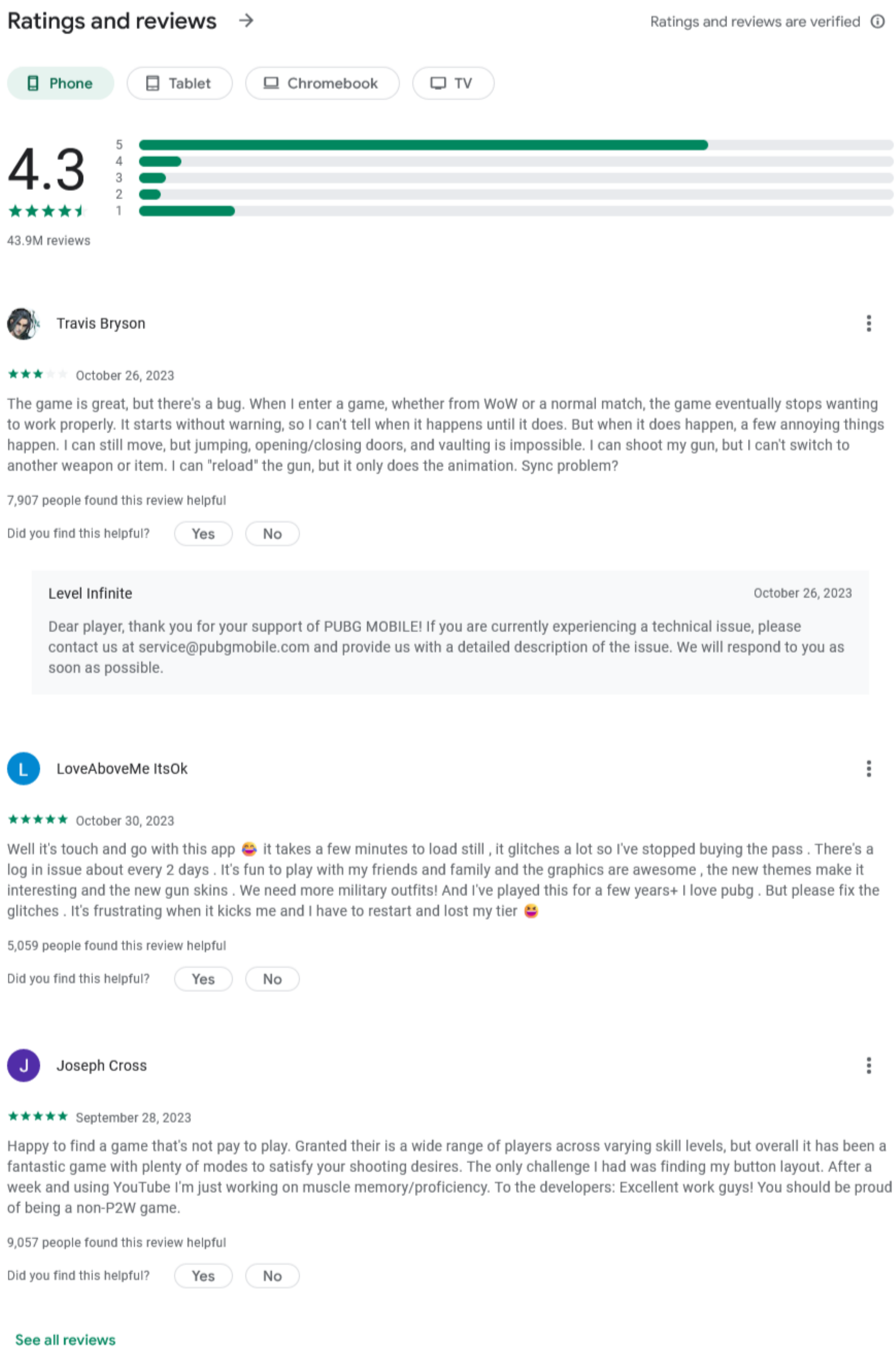


Figure 4.2: Steam Review Snapshot for PUBG Mobile.

feature that promotes the most pertinent feedback, aiding others in their decision-making process. "PUBG Mobile" showcases its multi-platform capability, with review filters for various devices, signifying its adaptability and inclusive approach that caters to a diverse gaming demographic. This integrative review system not only enriches the feedback loop but also signifies the game's commitment to addressing its diverse user base's needs.

4.3.2 Collection Method on Google Play

Utilizing the GooglePlay Developer API

The Google Play Developer API offers a robust interface for accessing app data, including reviews, from Google Play. This API is particularly useful for developers and researchers who need to programmatically interact with Google Play for data retrieval. It allows for the fetching of reviews from specific apps, providing insights into user feedback and app performance.

When collecting data on a large scale, such as for multiple apps, the API's ability to handle bulk requests is invaluable. By specifying the app IDs of interest, users can retrieve reviews for numerous apps in a single operation. This approach is efficient and minimizes the risk of exceeding API rate limits. The API also offers flexibility in terms of data customization. Users can specify parameters such as language, review type (e.g., positive, negative, or all), and time range, enabling targeted data collection tailored to specific research needs.

However, it's important to note that while the API facilitates comprehensive data collection, there may be limitations in terms of the total number of reviews retrievable for each app. Therefore, it is advisable to plan data collection strategically, possibly focusing on specific time frames or review types to ensure a representative dataset.

Direct Interaction with Google Play's Review API

Google Play's official API endpoint for retrieving app reviews is structured as follows: This API is highly configurable, allowing for the sorting of reviews, language specification, and filtering based on various criteria such as rating, date, and helpfulness. The output is typically in JSON format, providing detailed information about each review, including the reviewer's ID, review text, rating, timestamp, and more.

For our data collection, we utilized a specific API[2] to gather approximately six million English-language reviews from Google Play, up to March 2023. To refine this dataset for our analysis, we selectively excluded certain columns that were deemed non-essential, focusing on the most relevant data. The following table provides a detailed description of each column that was retained in our final dataset.

- **reviewId**: A unique identifier assigned to each review, ensuring the distinctiveness and traceability of the feedback.
- **userName**: The name of the user who posted the review.
- **userImage**: The URL or data representing the reviewer's profile image.
- **content**: The full text content of the user's review.
- **score**: The star rating (from 1 to 5) given by the user to the app.
- **thumbsUpCount**: The number of "likes" or "thumbs up" that the review has received from other users.
- **reviewCreatedVersion**: The specific version of the app that was being used when the review was written.
- **at**: The date and time at which the review was posted.
- **replyContent**: The content of any response provided by the app's developers to the review.
- **repliedAt**: The date and time at which the developer's reply was posted.

4.4 Text Preprocessing

4.4.1 Text Preprocessing

In our quest to unearth meaningful insights from the vast trove of collected reviews using LDA topic modeling, it became imperative to rigorously preprocess the textual

[2]https://pypi.org/project/google-play-scraper/

data. Our approach was significantly enriched by the capabilities of the Natural Language Toolkit (NLTK), a revered platform in the realm of natural language processing, known for its prowess in handling human language data within Python environments.

Noise Removal

The integrity and reliability of data play a pivotal role in ensuring the accuracy of topic modeling. Our noise removal process consisted of several meticulously planned steps:

- HTML Tags and Whitespace Removal: We began by purging the data of HTML tags, which often carry no semantic meaning for our analytical pursuits. Additionally, extraneous whitespace, which could skew our analysis or add unnecessary clutter, was removed.

- Language Detection and Filtering: With game reviews sourced globally, there's an inherent diversity in the languages used. It was imperative to ensure uniformity in language for our dataset. Utilizing advanced language detection tools, we identified and retained reviews in our target language while filtering out the rest. This ensured that the subsequent analytical tools operated with optimal accuracy.

- Spelling Correction: Player reviews, often written hastily or in the heat of the moment, are prone to typographical errors. These errors, if uncorrected, could lead to misinterpretation or skew our analysis. Harnessing automated spelling correction tools and dictionaries, we rectified these errors, preserving the integrity of the data.

- Special Character Deletion: Special characters, unless part of the game's jargon or semantically relevant, were removed. These characters can sometimes introduce noise or be misinterpreted by analytical algorithms.

- Elimination of Stop Words: Commonly used words like "and", "the", or "is" often add little to no value in topic modeling. We used the English stop word list to identify and eliminate these from our dataset, ensuring our algorithms focused on the more meaningful and relevant terms.

Through these steps, we meticulously cleansed the dataset, ensuring it was devoid of any elements that could compromise the accuracy or reliability of our subsequent analysis.

Normalization

Given the global reach of Steam's user base, our dataset was replete with diverse linguistic styles and idiosyncrasies. It was paramount to introduce a level of uniformity, and normalization was our first choice. A two-pronged approach was adopted: firstly, every piece of text was uniformly transformed to lowercase, ensuring that variations like "Game", "game", and "GAME" were treated as identical entities. Secondly, lemmatization was performed. This involved a rigorous morphological analysis of words, aided by NLTK's adept lemmatizer, ensuring that words were distilled to their canonical or dictionary forms. Such normalization is quintessential, as it amalgamates similar word manifestations under a singular, standardized term, reinforcing the fidelity of any ensuing analyses.

Tokenization

The last step of our preprocessing phase was tokenization, a technique wherein each review was meticulously broken down into individual words or terms. This transformed continuous streams of text into distinct, analyzable units known as tokens. By leveraging NLTK's state-of-the-art tokenization utilities, we ensured surgical precision in this decomposition. This foundational step not only played a pivotal role in identifying recurrent terms but also laid the groundwork for all subsequent preprocessing stages. The essence of tokenization lies in its potential to spotlight word frequencies, unveil term co-occurrences, and set the stage for deeper, more nuanced textual analyses.

4.5 Dataset Annotation

The cornerstone of our ABSA framework was the assembly of a robust dataset, for which we manually sifted through an extensive collection of esports game reviews, extracting 3,100 sentences that were rich in sentiment-laden aspects. Recognizing the nuanced nature of sentiment analysis, we engaged three annotators, each boasting substantial experience, to immerse themselves in this task. Their objective was clear: to dissect each sentence and identify explicit aspects, subsequently assigning appropriate sentiment polarities.

To maintain a uniform standard and high accuracy across annotations, we distributed the workload evenly among the annotators. Each was responsible for 1,000 sentences, encompassing a diverse range of reviews from three distinct esports games. Additionally, we introduced a set of 100 test sentences, a strategic move designed to serve as a common ground for evaluating inter-annotator agreement.

Annotation Strategy and Encoding

Central to our annotation process was the BIEOS encoding format, a structured tagging approach facilitating detailed sentiment analysis. The annotators were instructed to assign specific tags to words or phrases within the sentences: B-{POS, NEU, NEG}, I-{POS, NEU, NEG}, E-{POS, NEU, NEG}, O-{POS, NEU, NEG}, and S-{POS, NEU, NEG}. These tags were indicative of the position (Beginning, Inside, End, Outside) of the identified aspects in the sentiment expression, with a special "S" tag reserved for instances where a single word formed the entire aspect. The polarities POS, NEU, and NEG corresponded to Positive, Neutral, and Negative sentiments, respectively. For illustrative clarity, Table 4.1 showcases a sentence, exemplifying our annotation approach with the BIEOS encoding labels.

Table 4.1: An annotated sentence with labels in BIEOS encoding.

Words:	Toxic	community	and	smurfs	with	a	trashy	anti	cheat	system	.
Labels:	O	S-NEG	O	S-NEG	O	O	O	B-NEG	I-NEG	E-NEG	O

Inter-Annotator Agreement (IAA)

In the realm of data annotation, especially in tasks that involve subjective judgment such as sentiment analysis, establishing a consistent standard among different annotators is paramount. This consistency, often quantified as Inter-Annotator Agreement (IAA), is a statistical measure that evaluates the degree of agreement among annotators to ensure that the data is not only reliable but also representative of real-world scenarios and free from individual bias. One of the most widely recognized methods to calculate IAA is Cohen's kappa (κ) coefficient [58], a metric that accounts for the possibility of the agreement occurring by chance. Cohen's kappa is particularly advantageous because it provides a normalized score between -1 and 1, where a score of 1 represents perfect agreement, zero indicates no agreement better than chance, and negative values imply agreement less than chance, highlighting a systematic disagreement between annotators.

The formula for Cohen's kappa is:

$$\kappa = \frac{p_o - p_e}{1 - p_e} \tag{4.1}$$

where p_o is the proportion of times the annotators agree (observed agreement), and p_e is the proportion of times the annotators would be expected to agree by chance (expected agreement).

In our study, we calculated the IAA using Cohen's kappa for each pair of annotators and then averaged the results to accommodate the trio of annotators involved. This approach, though not as comprehensive as Fleiss' kappa, provided a pragmatic and straightforward assessment of annotator consistency. The results, detailed in Table 4.2, demonstrated a high level of agreement, indicative of the reliability of the annotations across different individuals. This high degree of IAA is crucial as it underscores the quality of the training data, which subsequently impacts the performance of the machine learning models trained on this data. By ensuring rigorous IAA, future researchers can confidently use our annotated data, knowing it accurately reflects the consensus among experts and serves as a solid foundation for the subsequent phases of the machine learning or data analysis project.

Table 4.2: Inter-annotator agreement for three annotators.

Annotator	Dota2	PUBG	CS:GO
A1 & A2	0.8784	0.9087	0.9309
A2 & A3	0.8669	0.8932	0.8840
A1 & A3	0.8365	0.8741	0.8512
#Average	0.8606	0.8920	0.8887

4.6 Ethical Considerations

In our comprehensive data collection encompassing various platforms, including Steam and Google Play, we have consistently adhered to the highest ethical standards. While collecting reviews from these platforms, we were mindful to exclude any information that could potentially lead to the identification of individual users. Our process strictly followed the terms of service of both Steam and Google Play, ensuring that the integrity of our research was maintained by concentrating on the content of the reviews rather than the identities of the reviewers.

For Steam, our data gathering methods were in strict compliance with the guidelines provided in the platform's robots.txt[3] file, ensuring that our automated tools did not intrude upon or scrape data from areas that were restricted. This approach[4] was mirrored in our handling of data from Google Play, where we similarly avoided collecting or storing any personal user information that could be used to trace back to individual users.

By respecting the privacy settings and preferences of users across both platforms, our methodology ensured that the collected data could not be retroactively associated with any particular individual. This precaution was fundamental to preserving the privacy and confidentiality of the users whose reviews contribute valuable insights into the esports gaming landscape.

To summarize, our data collection methodology, while sourcing esports game reviews from both the Steam and Google Play platforms, was rooted in ethical diligence. This careful approach ensured the privacy of users was protected, and the integrity and credibility of our research were upheld.

[3] https://store.steampowered.com/robots.txt
[4] https://play.google.com/robots.txt

Chapter 5

Methodology

5.1 Workflow

This research hinges on a structured workflow, essential for navigating the extensive and complex nature of player feedback in esports. The workflow is methodologically developed to ensure a systematic, unbiased, and thorough analysis, comprising distinct, interconnected stages. This workflow, depicted in Fig. 5.1, is a schematic representation of the analytical journey, demarcating several phases that contribute to the holistic understanding of player sentiment and topic prevalence.

5.1.1 Data Collection & Preprocessing

This phase stands as the cornerstone of the analytical journey, establishing the bedrock for the multifaceted evaluations that succeed it. This initial segment underscores its critical importance by embarking on a comprehensive assembly of data, drawing from a variety of sources to amass a wide-ranging collection of player feedback. Following this, the data is subjected to a stringent preprocessing sequence, an essential step that amplifies its fidelity and readies it for subsequent, detailed scrutiny. This refinement process is tripartite, encompassing Noise Removal, Normalization, and Tokenization, each of which is pivotal in fortifying the data's structure and integrity. Detailed in section 4.4.1, these collective measures are instrumental in fortifying the data's dependability, thereby ensuring that the conclusions drawn from future analyses are rooted in data that is not only pristine and uniform but also of superior quality.

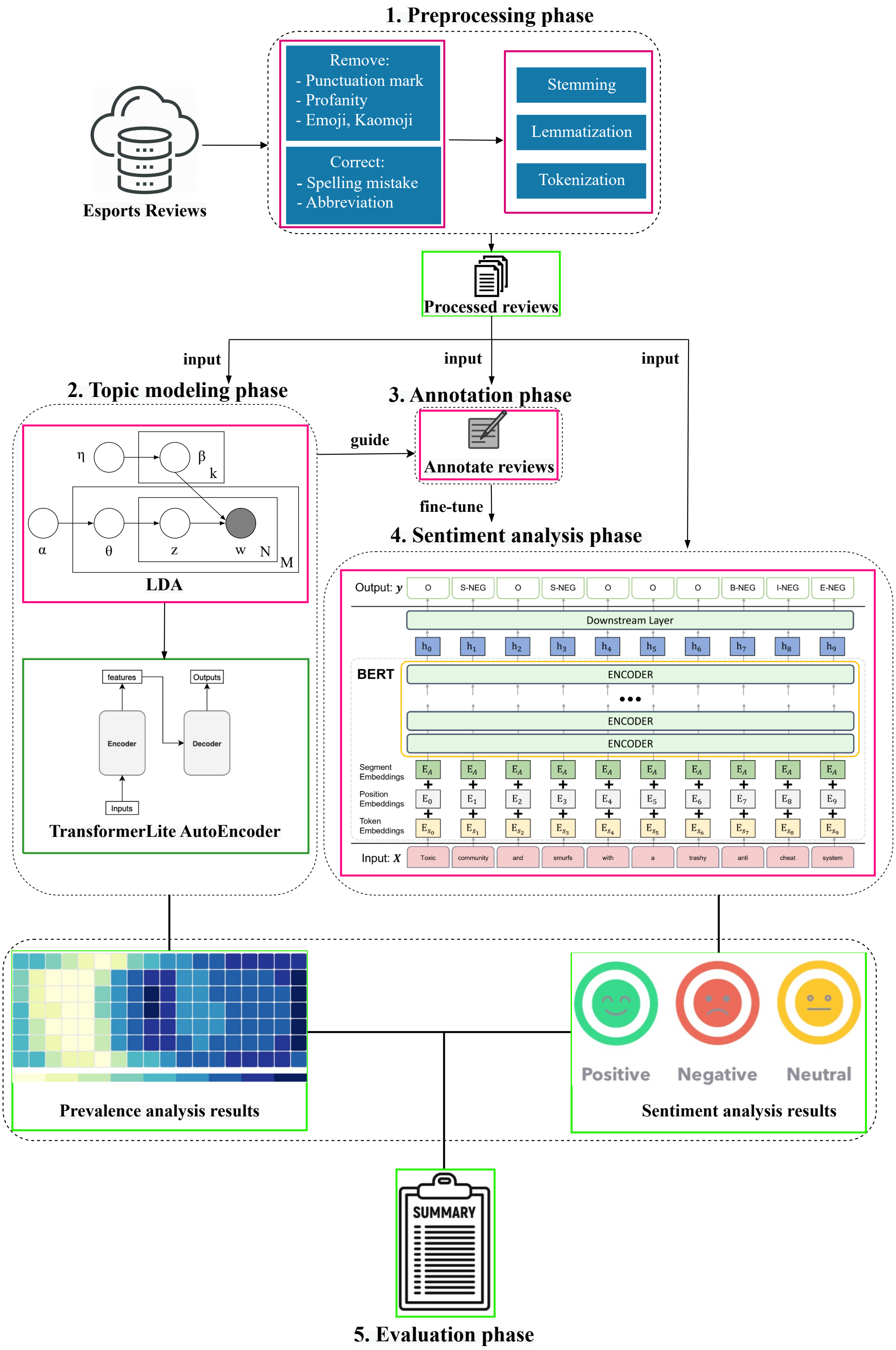


Figure 5.1: The workflow of the proposed framework.

5.1.2 Topic Modeling and Grouping

Initiating the process is the "Topic Modeling and Grouping" phase. Here, we delve into the realm of machine learning, utilizing sophisticated algorithms to parse through the enormity of unstructured data, thereby unraveling prevalent themes and topics resonating within the player feedback. The crux of this phase is the application of the Latent Dirichlet Allocation (LDA) model, a statistical approach that discerns underlying thematic structures and accordingly clusters the data. This phase is instrumental in transmuting raw, textual information into a structured format, ripe for further analysis. It's a convergence point where statistical modeling meets linguistic expertise, ensuring the thematic groups are coherent, representative, and set the stage for subsequent in-depth analysis.

Building upon the above phase with LDA, we further refine our analysis of *LDA-TransformerLite* architecture in the section**??**. This segment synergizes LDA's thematic detections with BERT's sentence embeddings, optimized through game-specific hyperparameters. The integration culminates in a TransformerLite AutoEncoder, adept at reducing dimensionality while preserving the depth and context of the data. This enhanced approach leads to a more nuanced extraction of topics.

5.1.3 Topic Prevalence Analysis

Transitioning from the realm of topic discovery, the workflow advances into the "Topic Prevalence Analysis" phase. This segment is an investigative probe into the significance of each discovered topic, quantifying their presence within the community discourse. It's a dual-faceted approach that melds quantitative frequency analysis with qualitative contextual examination. This analytical stage is pivotal, shedding light on the player community's focal points, thereby guiding stakeholders on areas requiring attention or improvement. It points towards topics that are not just frequently mentioned but also those that stir significant sentiment within the community.

5.1.4 Topic Sentiment Analysis

Culminating the workflow is the "Topic Sentiment Analysis" phase, a deep dive into the emotional undercurrents of the player feedback corresponding to each topic. This phase harnesses the power of sentiment analysis tools, interpreting the complex spectrum of human emotions reflected in the feedback. This involves a ternary (Positive, Neutral and Negative) classification of sentiments. This final stage is a critical juncture where insights are not just derived but also given context, empowering decision-makers with data-driven, actionable insights.

5.2 Topic Modeling with LDA-TransformerLite

5.2.1 Introduction

Latent Dirichlet Allocation (LDA) is a pioneering machine learning approach specifically designed for topic modeling, which automatically identifies topics in text corpora. Its relevance in analyzing voluminous digital texts, such as online game reviews, is profound. LDA helps in discerning patterns and insights that are not overtly apparent, thereby revealing hidden thematic structures and contributing significantly to the understanding of large-scale player feedback. This technique is particularly pertinent in the gaming industry, where player reviews are rich with diverse sentiments, opinions, and experiences.

5.2.2 Preliminaries

The foundational premise of LDA is that documents are composed of a mix of various topics, and these topics essentially consist of a collection of words with certain probabilities associated with them. Here, a “document” refers to the individual unit of text analysis (e.g., a single game review), and a “corpus” denotes the entire collection of documents being analyzed.

Key terms used in LDA include:

- **Document**: A piece of text (e.g., a game review).
- **Corpus**: A collection of documents.
- **Term**: A distinct word or phrase within the text.
- **Topic**: A recurring theme represented by a collection of terms.

Besides, the statistical model behind LDA involves several parameters:

- **Alpha** (α): A parameter affecting topic sparsity in documents. A lower value results in documents being composed of fewer topics.
- **Beta** (β): A parameter influencing word sparsity in topics. A lower value means that topics are composed of fewer words.

LDA operates under the assumption that the words in each document are not random and that certain words are more likely to co-occur, forming topics. Thus, the generative process for each document in LDA is as follows:

1. Choose a distribution over topics from the Dirichlet distribution with the parameter α.
2. For each word in the document:
   (a) Choose a topic from the distribution over topics.
   (b) Select a word from the corresponding distribution over the vocabulary, given the topic, with probability influenced by β.

5.2.3 Generative Process of LDA

The generative process of Latent Dirichlet Allocation (LDA) is a sophisticated, probabilistic approach that models each document in a corpus as a mixture of various topics, and each word as attributable to one of these topics. This intricate procedure involves multiple computational layers, each integral to the model's ability to accurately deduce the hidden thematic structures within the documents. The process begins as follows:

for each document d_d in corpus D **do**
    Choose $\theta_d \sim \text{Dirichlet}(\alpha)$, where θ_d is the topic distribution for document d.
    for each word position w in document d_d **do**
        Choose a topic $z_w \sim \text{Multinomial}(\theta_d)$.
        Choose a word w_w from $p(w_w|z_w, \beta)$, a multinomial probability conditioned on the topic z_w and the prior β.

6: **end for**

7: **end for**

This can be further broken down into the following steps to facilitate understanding:

1. For each topic, generate a distribution over words.

2. For each document, create a distribution over topics.

3. For each word in the document:

   (a) Select a topic based on the document's distribution over topics.

   (b) Choose a word according to the topic's distribution over words.

The graphical representation of the LDA procedure is depicted in Fig. 5.2.

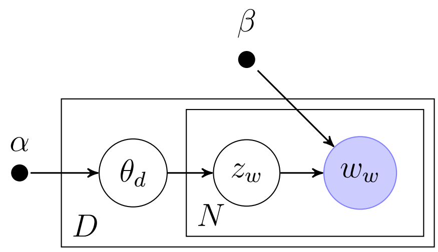


Figure 5.2: Latent Dirichlet Allocation model.

In mathematical terms, LDA seeks to infer the following distributions:

1. The term distribution $p(t|z = k) = \vec{\phi_k}$ for each topic k.

2. The topic distribution $p(z|doc = d) = \vec{\theta_d}$ for each document d.

This is represented by the joint distribution of the topic mixtures Θ, the set of topic assignments Z, the terms in the corpus W, and the topics Φ, given the hyperparameters α and β. The equation for this joint distribution is:

$$P(Z, \Theta, \Phi, W|\alpha, \beta) = \prod_{i=1}^{K} P(\phi_k|\beta) \prod_{d=1}^{M} P(\theta_d|\alpha) \prod_{n=1}^{Nd} P(z_{d,n})P(w_{d,n}|\phi_{z_{d,n}}) \tag{5.1}$$

The goal is to compute the posterior distribution:

$$P(Z, \Theta, \Phi | W, \alpha, \beta) = \frac{P(Z, \Theta, \Phi, W | \alpha, \beta)}{P(W | \alpha, \beta)} \tag{5.2}$$

However, the actual computation of the posterior distribution of the hidden variables given the observed variables is not straightforward due to the complexity of the model's structure. Instead of directly computing this distribution, LDA employs approximation methods. These strategies aim to optimize the likelihood of the observed documents based on the parameters, circumventing the direct calculation of the complex denominator that requires summation over every potential configuration of Z, Θ, and Φ.

In the realm of online game reviews, LDA serves as a powerful tool for unveiling hidden thematic structures and player sentiments. By deciphering the topics that players commonly discuss, stakeholders can gain nuanced insights that drive data-informed decisions, enhancing gaming experiences and fostering a more engaged player community.

5.2.4 The LDA-TransformerLite Architecture

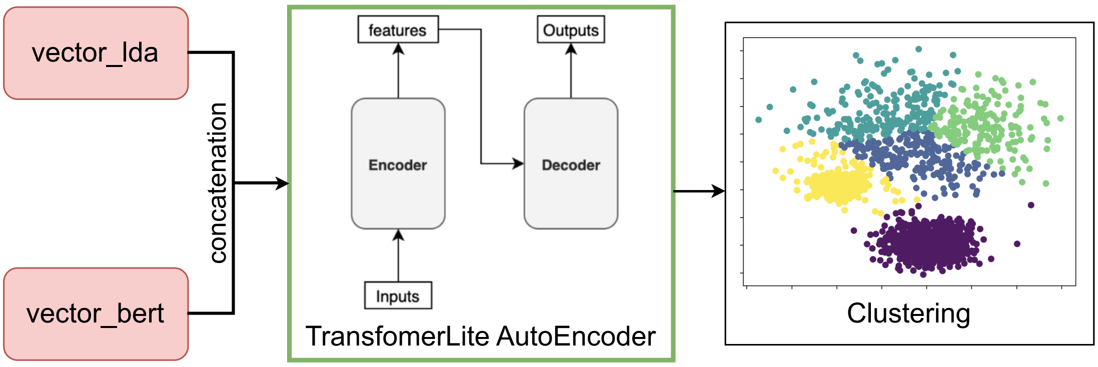


Figure 5.3: The workflow of the proposed framework.

Building on the foundation laid by the introduction of LDA for topic inference, the model takes a significant leap forward with the integration of sentence embeddings obtained from BERT as shown in Fig. 5.3. This approach entails two key steps:

1. We combined the probabilistic topic assignment vector from LDA and the sentence embedding vector from BERT into a single composite vector. To ensure a balanced

integration of information from both sources, we utilized a weight hyperparameter γ, set to twice the value of k, which represents the number of topics. This approach allows us to adjust the influence of each vector, achieving an optimal mix of semantic and contextual information tailored to different esports games.

2. Recognizing the challenges presented by the high-dimensional nature of the concatenated vector, where information is often sparse and interrelated, the model employed a TransformerLite AutoEncoder for dimensionality reduction. This advanced architecture harnesses the power of self-attention mechanisms, a hallmark of transformer models, to deliver a more nuanced understanding of data sequences and structures. The TransformerLite AutoEncoder adeptly maps the high-dimensional input to a lower-dimensional latent space, concentrating the data's essence while preserving its contextual integrity.

We lastly apply clustering techniques on the model and the result is a more refined latent space representation. This enhancement markedly improves the model's ability to distill contextual topics from the clusters, leading to a topic modeling process that is not only more contextually aware but also more nuanced in its understanding of the textual data.

The following Table 5.1 presents the detailed architecture of the TransformerLite AutoEncoder implemented in the model. It provides an overview of each layer, including its type, output shape, and the number of parameters. This table illustrates the complexity and intricacy of the TransformerLite AutoEncoder, showcasing how it effectively processes and transforms the high-dimensional input data.

5.2.5 Topic Coherence

Topic coherence measures the degree of semantic similarity between high scoring words in a topic. These scores provide a window into the model's quality, as topics with higher coherence scores are typically more interpretable and make more sense to human evaluators. In the context of u_mass coherence [59], the score is calculated based on the co-occurrence of the most prominent words within each topic, considering the entire corpus.

Table 5.1: TransformerLite AutoEncoder Architecture Description

Layer (type)	Output Shape	Param #
input_1 (InputLayer)	(None, 768 + k)	0
dense (Dense)	(None, 256)	200,704
tf.expand_dims (TFOpLambda)	(None, 1, 256)	0
multi_head_attention (MultiHeadAttention)	(None, 1, 256)	1,577,728
dropout (Dropout)	(None, 1, 256)	0
tf.__operators__.add (TFOpLambda)	(None, 1, 256)	0
layer_normalization (LayerNormalization)	(None, 1, 256)	512
dense_1 (Dense)	(None, 1, 128)	32,896
dense_2 (Dense)	(None, 1, 256)	33,024
dropout_1 (Dropout)	(None, 1, 256)	0
tf.__operators__.add_1 (TFOpLambda)	(None, 1, 256)	0
layer_normalization_1 (LayerNormalization)	(None, 1, 256)	512
tf.compat.v1.squeeze (TFOpLambda)	multiple	0
dense_3 (Dense)	(None, 783)	201,231

The process begins by identifying the most frequent M words in a topic, represented as $V^{(t)} = \left\{v_1^{(t)}, \ldots, v_M^{(t)}\right\}$. Each word type v in $V^{(t)}$ is assessed based on its document frequency, $D(v)$, which counts the number of documents containing at least one instance of v. This frequency reflects the prevalence of each word type in the corpus. Furthermore, the co-document frequency, $D(v, v')$, measures the number of documents in which two word types, v and v', co-occur, highlighting their semantic relationship. The coherence for topic t is then calculated using the formula:

$$C\left(t; V^{(t)}\right) = \sum_{m=2}^{M} \sum_{l=1}^{m-1} \log \frac{D\left(v_m^{(t)}, v_l^{(t)}\right) + 1}{D\left(v_l^{(t)}\right)} \tag{5.3}$$

The u_mass coherence score is designed such that the co-occurrence of frequently appearing words within topics contributes to a higher (less negative) coherence score. Essentially, if the top words of a topic frequently appear together in documents, the topic is deemed to be more coherent (makes more sense), thereby receiving a higher score. Since u_mass coherence scores are generally negative, the optimal point is often the "lowest negative" score, indicating the best balance between topic specificity and interoperability.

Yet, it's critical to perform manual verification of the findings, as authenticating the outputs of unsupervised machine learning processes is consistently complex. Reflecting on our experimental data, there were scenarios where the global minimum surged beyond 50, while the second-lowest global minimum registered under 20. In such instances, we opted for the local minimum as the definitive choice for our optimal number of topic assessments.

5.3 Topic Prevalence Analysis

In this research phase, we delve deeply into the realm of topic prevalence, utilizing a robust analytical approach that integrates the topic words inspired by Lin et al.'s work [45], as identified from the initial topic modeling stage, with the expansive corpus of player reviews. This method allows for a comprehensive exploration into the frequency and significance of specific topics within the esports gaming community, grounding the analysis in real-world data and ensuring its relevance and applicability.

Our approach to determining topic prevalence is encapsulated in a precise algorithm, represented by the formula:

$$\text{Topic Prevalence} = \frac{\text{\# of reviews which contain the specific topic}}{\text{\# of total reviews}} \tag{5.4}$$

This formula serves as the backbone of our analysis, providing a clear and quantifiable means of assessing the extent to which certain topics permeate the player discourse. The integrity of this process is upheld through the careful selection of keywords for the frequency analysis, avoiding those that could introduce ambiguity or statistical bias due to multiple meanings or generic usage.

Through this multifaceted methodology, our Topic Prevalence Analysis not only quantifies the presence of various topics in player reviews but also provides profound insights into the narratives and sentiments that underpin them. This phase is instrumental in identifying the focal points of player discourse, offering invaluable guidance for stakeholders in the esports industry on areas that require attention, improvement, or further exploration.

5.4 Sentiment Analysis

5.4.1 Introduction

Sentiment analysis operates at the intersection of artificial intelligence and linguistics, offering profound insights into public opinion by analyzing emotional undertones in text. Especially relevant on online platforms, sentiment analysis deciphers the complex layers of human emotion within digital communication. In online gaming, understanding player sentiment goes beyond mere likes or dislikes, delving into nuanced perceptions and attitudes towards different game elements. This analysis is crucial in shaping player-centric gaming experiences, informing developers about areas needing attention, and highlighting features that resonate well with the audience.

5.4.2 Preliminaries

Aspect-based Sentiment Analysis

Aspect-based Sentiment Analysis (ABSA) is a refined level of sentiment analysis that doesn't just seek to understand the overall sentiment of a text but delves deeper to extract sentiments about specific aspects or components within the subject matter. This approach is crucial, especially in scenarios where the general sentiment doesn't reflect the entire story, and opinions about individual aspects are mixed.

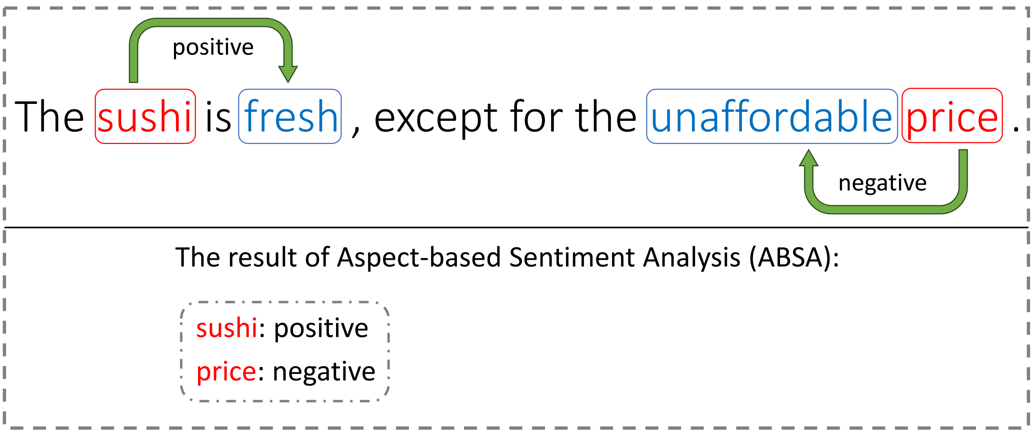


Figure 5.4: A representation of Aspect-based Sentiment Analysis.

To illustrate, let's consider a practical example encapsulated in Fig. 5.4. Take the sentence: "The sushi is fresh, except for the unaffordable price." In this instance, ABSA doesn't stop at identifying the overall positive sentiment but goes further to discern that the positivity is attributed specifically to the "sushi". Conversely, it recognizes a negative sentiment concerning the "price" aspect, all within the same sentence. This nuanced analysis is pivotal, particularly in domains like product reviews, where consumers may have distinct opinions on different attributes of a product or service. In Fig. 5.4, the aspect terms are highlighted in red, indicating the subjects being discussed, and the sentiment polarities are marked in blue, representing the emotional tone (positive, negative) associated with each aspect.

This method's strength lies in its ability to provide a more granular analysis, offering detailed insights into specific strengths and weaknesses, rather than a broad-brushed sentiment evaluation. Such detailed sentiment dissection is invaluable for businesses and service providers to pinpoint areas for improvement and recognize strengths, thereby aligning closely with customer satisfaction and preferences.

The sentiment analysis of player reviews offers a window into the emotional undertones associated with each identified topic. By applying sentiment analysis algorithms to the corpus of reviews, we have been able to quantify the positivity, negativity, and neutrality in player feedback, thus gaining a deeper understanding of player satisfaction and areas of concern.

Pre-BERT Architectures

Before the advent of BERT, several NLP architectures were explored to enhance machine understanding of human language. Early models relied on hardcoded language rules, limiting their adaptability. Statistical models, like Bag of Words (BoW) and TF-IDF, offered more flexibility but lacked in understanding nuanced context. The advent of machine learning introduced models like Support Vector Machines (SVM) and Random Forest, learning to predict sentiment from labeled data offering better context understanding but still lacking in capturing language nuances fully. The real breakthrough came with sequence models like Recurrent Neural Networks (RNN) and their more advanced version, Long Short-Term Memory networks (LSTM), designed to capture sequential information, making them suitable for NLP tasks.

LSTM As the first neural network widely used in natural language, the structure of a recurrent neural network (RNN) naturally fits the characteristics of the sequence model. However, The pure RNN is difficult to deal with long-term dependence due to the vanishing/exploding gradient. As the most popular variant of RNN, Long Short-Term Memory [60] introduced the concept of "Gate" to almost perfectly solve the vanishing/exploding gradient problem in RNN. LSTM allows the weights to be changed at different times and allows the network to forget that it has accumulated information. In this paper, we used LSTM as one of the downstream layers. The computational formula of the hidden representation h_t and output o_t at the t-th time step is defined as:

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{5.5}$$

$$i_t = \sigma\left(W_i \cdot [h_{t-1}, x_t] + b_i\right) \tag{5.6}$$

$$\tilde{c}_t = \tanh\left(W_c \cdot [h_{t-1}, x_t] + b_c\right) \tag{5.7}$$

$$c_t = f_t * c_{t-1} + i_t * \tilde{c}_t \tag{5.8}$$

$$o_t = \sigma\left(W_o \left[h_{t-1}, x_t\right] + b_o\right) \tag{5.9}$$

$$h_t = o_t * \tanh\left(c_t\right) \tag{5.10}$$

where σ is the sigmoid activation function, and tanh is the tanh activation function. x_t is the input vector and h_t is the hidden state vector storing all state information until t-th time; c_{t-1} and c_t are the previous and current cell states at time t, respectively; W_f, W_i, W_c, W_o are the weight matrices for hidden state h_t; b_f, b_i, b_c, b_o denote the bias vectors. Formulas 5.5, 5.6, 5.7-5.8, 5.9 represent the four stages of LSTM: forget gate, input gate, cell state, and output gate, respectively. To conclude, the forget gate determines which relevant information from previous steps is needed. The input gate determines which relevant information can be added from the current step, and the output gate finalizes the next hidden state. The most important cell state gives the model long-term memory capability to store and load global information from previous steps.

However, these models faced challenges with long-term dependency information, leading to the development of the Transformer model, which set the stage for BERT.

Despite the advancements brought by sequence models like RNNs and LSTMs, and even the revolutionary context-awareness of the Transformer model, the NLP community still faced significant challenges.

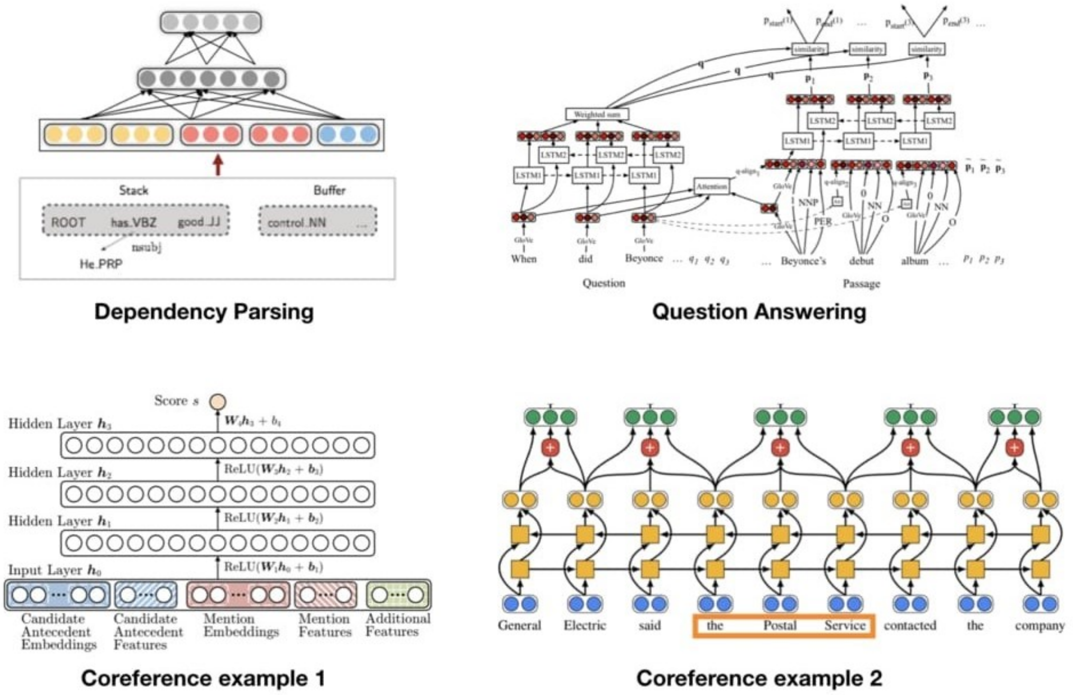


Figure 5.5: Various model architectures for different NLP tasks on Pre-BERT era.

Before BERT, the NLP landscape was populated with a variety of models, each tailored to specific tasks, as demonstrated in Fig. 5.5. Different NLP tasks often required unique models, and the process of designing these specialized models was labor-intensive, time-consuming, and required substantial computing resources. Each new task potentially needed a new model, making the field of NLP a costly endeavor both in terms of human labor and computational power.

Pretraining and Fine-Tuning in BERT

The advent of BERT revolutionized this model-specific landscape. BERT exemplifies a two-stage migratory learning approach that has surged in popularity within the NLP community. The genius of BERT lies in its pretraining stage, where a Transformer Encoder is used alongside a vast corpus of text and two pre-training objectives, creating a versatile model with a fundamental "understanding" of language semantics. This model doesn't just learn the structure of language; it grasps the nuance and context, attributes essential for complex NLP tasks. What follows is the fine-tuning stage, where this pre-trained model serves as a foundational architecture for various downstream NLP tasks. Whether for feature extraction or task-specific tuning, the same BERT model can be adapted. Fig. 5.6 highlights this uniformity, which significantly reduces the architectural engineering costs and resource expenditures traditionally associated with NLP tasks.

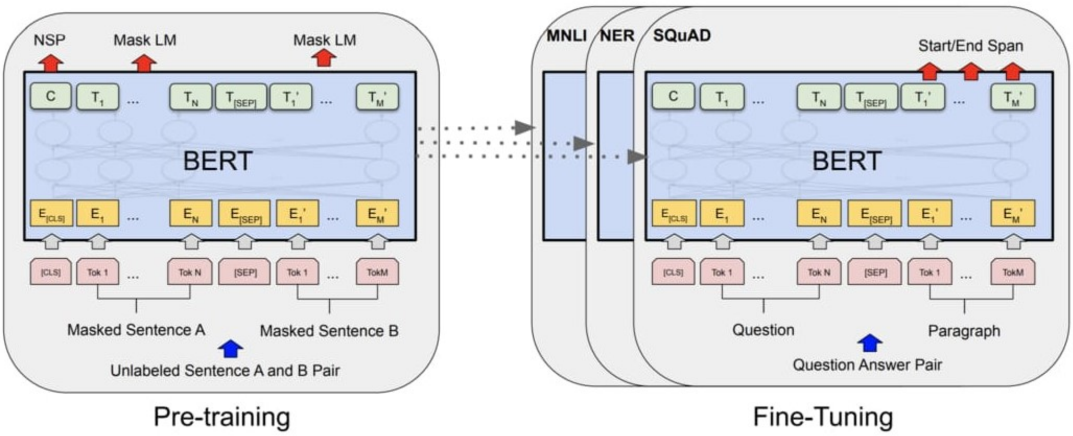


Figure 5.6: Two phases transfer learning in BERT model.

One of the pivotal concepts in this methodology is the symmetry between the pre-training and fine-tuning steps when using BERT. The architecture, pre-trained on a broad scale, is so versatile that it can be fine-tuned for multiple NLP tasks without substantial changes. This approach has streamlined the entire field of NLP, allowing researchers and engineers to bypass the traditionally high costs of designing and testing new models for each unique task.

By mastering BERT's application, one can efficiently proceed with numerous NLP tasks, harnessing the model's capacity to understand and generate language in a way that mirrors human cognition, all while saving on the time, labor, and costs previously deemed indispensable.

5.4.3 Sentiment Analysis Algorithm

After obtaining the BERT representations, we added a neural layer as the last layer to fine-tune the BERT model with different designs for our E2E-ABSA task as illustrated in Fig. 5.7. The layers we investigated included the linear layer as the baseline, and the transformer layer.

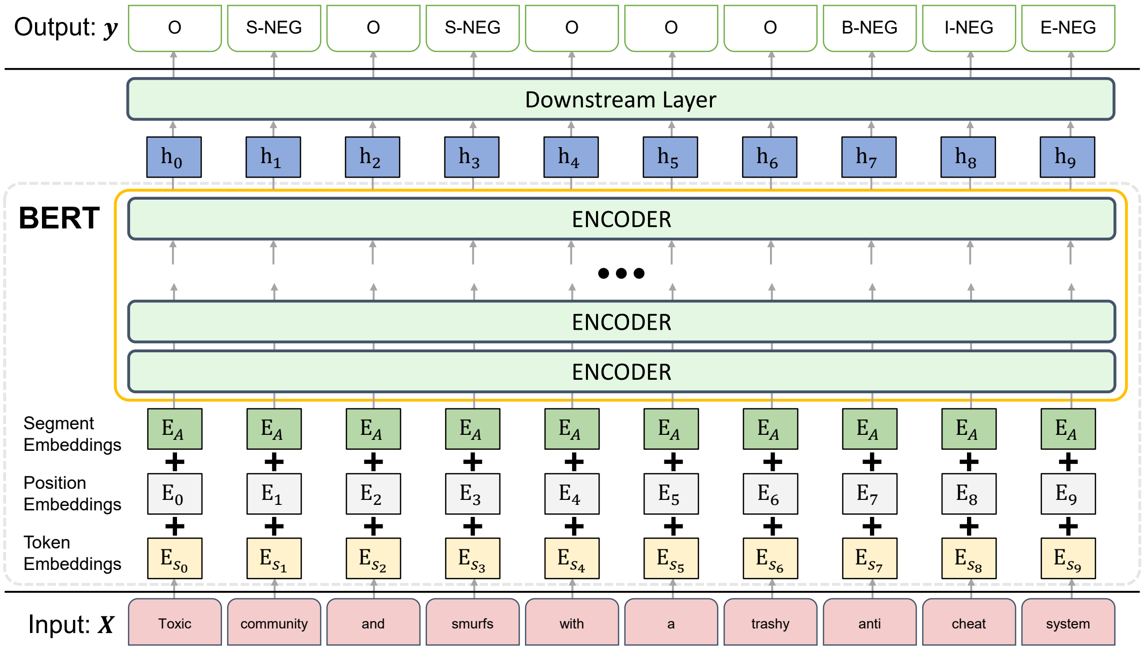


Figure 5.7: Overall of the E2E-ABSA model architecture.

BERT as Embedding Layer

BERT is a powerful language representation model developed by Google. It leverages masked language models to pre-train deep bidirectional representations, achieving impressive performance on sentence and token-level tasks. Given BERT's exceptional feature capture capabilities, we used it as the embedding layer to extract sentence information. Specifically, we utilized the pre-trained "bert-base-uncased model" [1] with default parameter settings, such as num_hidden_layers (number of hidden layers in the Transformer encoder) set to 768 and the max_position_embeddings (the maximum sequence length that this model can be used with) set to 512. These parameters can be adjusted to suit the size of the model.

To ensure that the review text adhered to BERT's input-size limitations, we checked the token size of each review and split any overlength reviews into several paragraphs. We also manually reviewed all feedback and removed any instances of

[1] https://github.com/huggingface/transformers

ASCII art steam reviews to obtain a cleaner dataset for input. For a given sentence $X = (x_1, x_2, \ldots, x_t)$, where $1 \leq t \leq 512$ is the length of the sentence, the embedding space is represented by vectors encapsulating the meaning of each word, with similar words having closer vector values. BERT's input embeddings are packaged as $A = (a_1, a_2, \ldots, a_t)$, which is the sum of the token embeddings, the segmentation embeddings, and the position embeddings. The remainder of the implementation is virtually identical to that of the original BERT, and we refer to the transformer layer without delving into an in-depth description of the model architecture.

Liner Layer

The linear layer containing a single softmax activation function calculates the token-level predictions based on the token representations obtained from the previous layer:

$$P(y_t|x_t) = \text{softmax}\left(Wx_t + b\right) \tag{5.11}$$

where W and b are the liner layer's learnable matrices.

Transformer (TFM)

Once we obtained the BERT representations, we added a transformer layer as the final layer to fine-tune the BERT model with various designs for the E2E-ABSA task. The Transformer model, proposed by Vaswani et al. [61] provides a robust feature extractor while discarding traditional RNN architectures in NLP tasks. Additionally, since BERT is fundamentally made up of the encoder of the Transformer model, we used a Transformer layer with the same architecture as the BERT encoder in this study. The TFM's computational process is straightforward and can be broken down into several steps, as shown below:

$$Q_i, K_i, V_i = \left(W_i{}^Q, W_i{}^K, W_i{}^V\right) A \tag{5.12}$$

$$\mathrm{Att}_i(Q, K, V) = \mathrm{softmax}\left(\frac{Q_i K_i{}^T}{\sqrt{d_k}}\right) V_i \tag{5.13}$$

$$\text{Self-Att}_{MH}(Q, K, V) = \mathrm{Concat}\left(\mathrm{Att}_i\right) W^O \tag{5.14}$$

$$\hat{A} = \mathrm{LN}\left(A + \text{Self-Att}_{MH}(Q, K, V)\right) \tag{5.15}$$

$$y = \mathrm{LN}\left(\hat{A} + \mathrm{FFN}\left(\hat{A}\right)\right) \tag{5.16}$$

In the above equations, Q, K, and V represent the Query, Key, and Value vectors that are generated from the input vector. The feed-forward network (FFN) consists of a simple fully-connected neural network that uses ReLU as the activation function [61]. Formulas 5.12-5.14 illustrate the computational process of the multi-head self-attention mechanism with i heads, while Formulas 5.15-5.16 show the residual connection step [62] and layer normalization. Finally, a linear layer with softmax activation is attached to the output of the TFM layer to obtain the final prediction.

5.4.4 Evaluation Method

The efficacy of our sentiment classification approach was gauged using a standard three-class confusion matrix, delineated in Table 5.2, which provides a detailed account of the classification outcomes. The predictive accuracy is categorized into four distinct outcomes: True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN). Here, the terms True (T) and False (F) reflect the accuracy of the prediction, while Positive (P) and Negative (N) denote whether the sentiment is positive or negative.

In our analysis, we assigned three distinct sentiment labels, denoted as $\mathbf{T} = \mathrm{POS}, \mathrm{NEU}, \mathrm{NEG}$, corresponding to Positive, Neutral, and Negative sentiments. The computation of TP, FP, TN, and FN for each sentiment category is articulated through Equations 5.17, 5.18, and 5.19.

Table 5.2: Confusion matrix for three sentiment labels.

Confusion Matrix		Actual		
		POS	NEU	NEG
Predicted	POS	a	d	g
	NEU	b	e	h
	NEG	c	f	i

- Calculations for the POS label:

$$\begin{aligned} TP_{POS} &= a \quad & TN_{POS} &= e + f + h + i \\ FP_{POS} &= d + g \quad & FN_{POS} &= b + c \end{aligned} \tag{5.17}$$

- Calculations for the NEU label:

$$\begin{aligned} TP_{NEU} &= e \quad & TN_{NEU} &= a + c + g + i \\ FP_{NEU} &= b + h \quad & FN_{NEU} &= d + f \end{aligned} \tag{5.18}$$

- Calculations for the NEG label:

$$\begin{aligned} TP_{NEG} &= i \quad & TN_{NEG} &= a + b + d + e \\ FP_{NEG} &= c + f \quad & FN_{NEG} &= g + h \end{aligned} \tag{5.19}$$

Micro-Level Evaluation For the *Micro* F_1 assessment, we aggregated the total counts of TP, FN, and FP across all sentiment categories to ascertain the universal *Precision* and *Recall*:

$$\begin{aligned} Precision_{mi} &= \frac{\sum_{t \in T} TP_t}{\sum_{t \in T} TP_t + \sum_{t \in T} FP_t} \\ Recall_{mi} &= \frac{\sum_{t \in T} TP_t}{\sum_{t \in T} TP_t + \sum_{t \in T} FN_t} \end{aligned} \tag{5.20}$$

Following the methodology used for $MacroF_1$, the $MicroF_1$ score is determined as:

$$Micro - F_1 = \frac{2 * Precision_{mi} * Recall_{mi}}{Pre_{mi} + Recall_{mi}} \tag{5.21}$$

Chapter 6

Dataset and Result Analysis

6.1 Experimental Dataset

After completing the preprocessing steps, our dataset was meticulously cleaned and readied for in-depth topic modeling. It was now primed for granular topic modeling. For more details on the results of our preprocessing, refer to Table 6.1, which shows the number of reviews categorized by sentiment and game.

Table 6.1: The properties of the processed datasets.

Game	PUBG	Dota2	CS:GO
#Recommended	189,776	397,090	726,186
#Not Recommended	117,353	61,024	99,912
#Total	307,129	458,114	826,098

Having our preprocessing, we embarked on an analysis phase. the temporal dimension of our reviews yielded intriguing insights. Notably, The longevity and persistent evolution of games make them especially susceptible to time-dependent feedback. In fact, Lin et al. [45] unveiled a fascinating correlation between a player's total playtime and their feedback, especially in the esports sector. Their research underscored the importance of playtime in gauging a player's sentiment towards a game.

Simultaneously, we directed our attention to the mobile esports segment, specifically focusing on the PUBG Mobile (PUBGm) dataset. Our efforts here were geared towards understanding a different facet of the esports universe, where player interactions and feedback might present unique patterns compared to PC games. Table 6.2

delineates the detailed distribution of the PUBGm dataset post-processing. This step was crucial to accurately capture the sentiments expressed by players, which would in turn inform our topic modeling.

Table 6.2: Detailed Properties for PUBGm dataset

Rating	Count	Percentage
1 star	925,769	14.87%
2 stars	151,392	2.43%
3 stars	225,403	3.62%
4 stars	373,452	6.00%
5 stars	4,549,245	73.08%
Total	6,225,261	100.00%

The distribution of ratings, reveals a significant skew towards positive feedback, with 73.1% of the ratings being 5 stars. This overwhelming majority underscores the generally favorable reception of PUBGm by its user base. However, a non-negligible portion of the ratings, amounting to 14.9%, were 1-star reviews, indicating a subset of players with critical perspectives on the game.

6.2 Topic Modeling and Grouping

In our investigation, we applied the Latent Dirichlet Allocation (LDA) method to unearth the underlying topics that pervade player reviews across four major esports games. Our methodology for determining the optimal number of topics was guided by coherence score, as delineated by Newman et al. [63]. The coherence scores were plotted against a range of potential topic numbers, from 3 to 50, as illustrated in Fig. 6.1. After careful analysis, we identified the most coherent separations at 16 topics for Dota2, CS:GO and PUBGm, and 15 for PUBG, where the coherence scores reached a plateau.

The visualized K-means results in Fig. 6.2 offer a revealing comparison between the conventional LDA topic modeling and our enhanced LDA-Transformer model. A consistent pattern observed across all games in our dataset is the distinct clarity and separation of topics achieved with the LDA-Transformer model. This is evident when comparing the density and overlap of clusters between the two approaches.

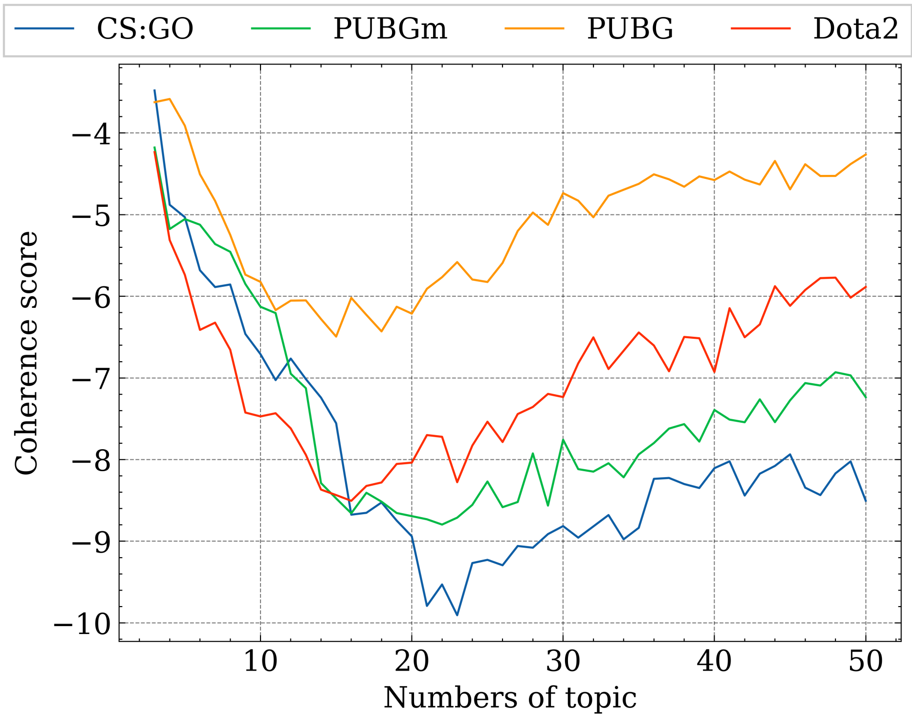
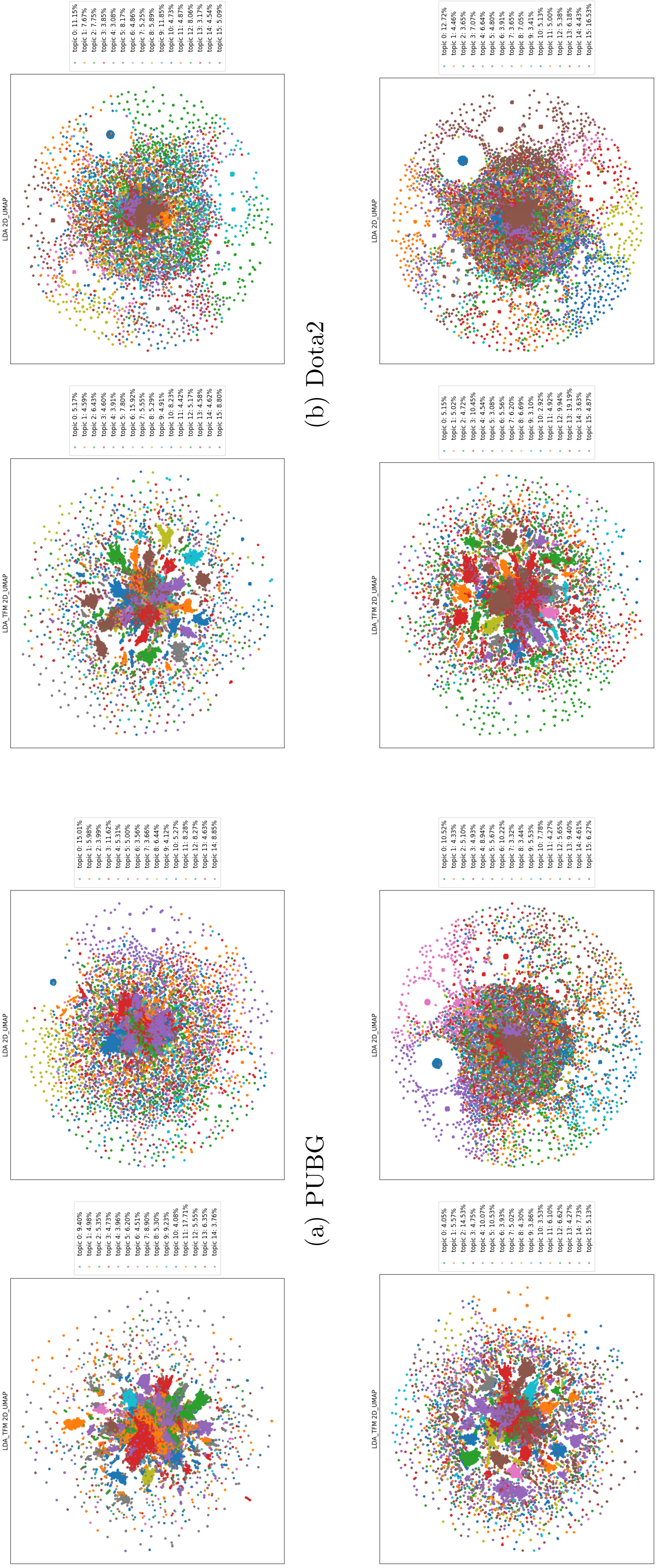


Figure 6.1: Coherence scores of four esports games.

The UMAP projections for the standard LDA show a high degree of topic aggregation, where different topics are often clustered closely together, making it challenging to discern distinct thematic groupings. This can be attributed to the conventional LDA's limitations in handling the high-dimensional space of complex textual data. In contrast, the LDA-Transformer model demonstrates a more articulate clustering pattern, with clear delineation between topics. The effectiveness of the Transformer-LiteAutoencoder in reducing dimensionality, while preserving the semantic and contextual nuances of the data, results in a more refined visual representation. The topics are not only more spread out across the 2D space but also form more defined clusters, indicating a better segregation of the semantic content of player feedback. These visual distinctions suggest that the LDA-TransformerLite model offers a more nuanced understanding of the underlying topics within the game reviews.

The WordCloud displays, as shown in Fig. 6.3, serve as a graphical elucidation of the topics distilled from the review corpus. In these visualizations, the scale of each term is proportionate to its frequency and relevance within a topic, offering a direct, visual interpretation of the data.



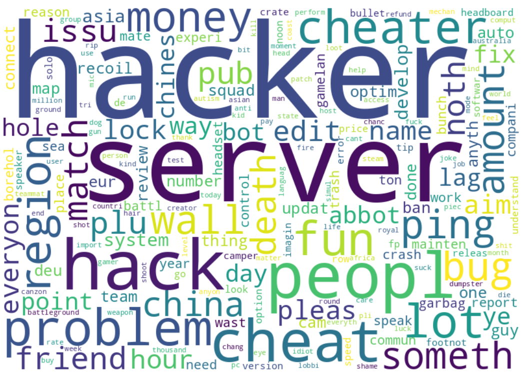
(a) PUBG

(b) Dota2

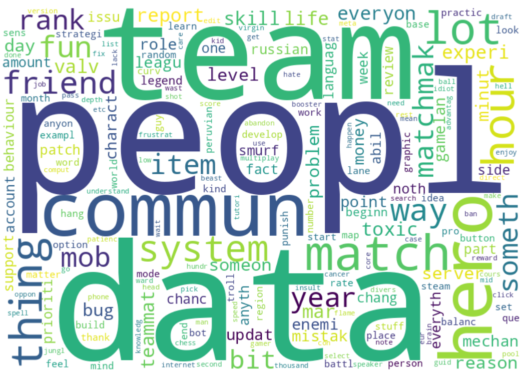
(c) CS:GO

(d) PUBGm

Figure 6.2: The comparisons of our proposed LDA-TransformerLite vs pure LDA by using K-means clustering.

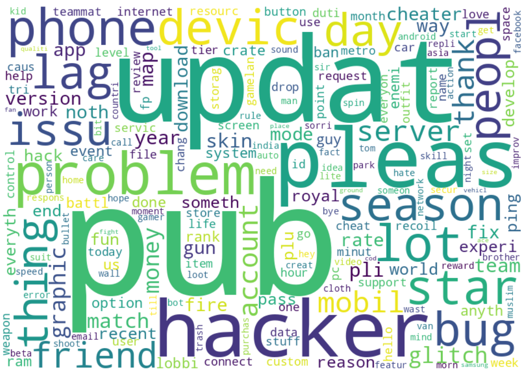


(a) PUBG



(b) Dota2

(c) CS:GO



(d) PUBGm

Figure 6.3: Four esports games topics Wordcloud.

To corroborate the integrity of the topics derived from the LDA model, we conducted a thorough examination of the WordClouds in conjunction with the original review texts. This cross-examination was imperative to sift through and exclude any potential noise, such as vague phrases or forum spam, which are prevalent in online discussions. The topics that emerged from this meticulous process are systematically enumerated in Table 6.3 and are segregated into two overarching classifications reflective of their content and implications.

Game-related Topics (GRT) The first group, labeled as GRT, encompasses topics with No. 1 to 6 in Table 6.3. These topics are the pillars of game design and are consistently referenced across reviews for PUBG, Dota2, CS:GO, and PUBGm. The "graphics" topic, for instance, includes discussions on visual fidelity and performance, which are critical for player immersion and satisfaction. The "character" topic captures conversations around character design, abilities, and customization options, which are vital for player identification and engagement with the game. The "map"

Table 6.3: Keyword examples and inferred topics of esports games.

(a) PUBG

No.	**Inferred topics**	**Keyword examples (PUBG)**
1	graphics	graphic*0.019, fps*0.004
2	character	player*0.018, skin*0.020
3	map	map*0.027, zone*0.007
4	optimization	crash*0.025, bug*0.052
5	update	update*0.029, release*0.011
6	gameplay	gameplay*0.006, battle*0.011
7	community	people*0.015, relationship*0.003
8	server	server*0.035, lag*0.007
9	region	region*0.027, china*0.033
10	price	money*0.028, lprice*0.004
11	matchmaking	match*0.010, hour*0.010
12	cheating	hacker*0.022, cheater*0.080
13	skill	gun*0.013, shoot*0.009
14	learning curve	learning*0.002, curve*0.003
15	teamwork	team*0.006, friend*0.018

(b) Dota2

No.	**Inferred topics**	**Keyword examples (Dota2)**
1	graphics	graphic*0.026, fps*0.002
2	character	hero*0.060, character*0.014
3	map	map*0.007, rock*0.002
4	optimization	bug*0.004, issue*0.002
5	update	update*0.044, version*0.010
6	gameplay	moba*0.065, multiplayer*0.006
7	community	community*0.048, people*0.038
8	server	server*0.024, internet*0.001
9	region	SEA*0.019, language*0.011
10	price	money*0.008, pay*0.011
11	matchmaking	match*0.015, matchmaking*0.006
12	cheating	smurf*0.005, hacker*0.002
13	skill	farm*0.007, lifesteal*0.002
14	learning curve	learning*0.019, noob*0.015
15	teamwork	team*0.036, teammate*0.028
16	ranking	rank*0.004, mmr*0.012

Table 6.3: Keyword examples and inferred topics of esports games (continued).

(c) CS:GO

No.	**Inferred topics**	**Keyword examples (CS:GO)**
1	graphics	graphic*0.008, fps*0.016
2	character	terrorist*0.008, character*0.002
3	map	map*0.021, zone*0.005
4	optimization	problem*0.021, issue*0.007
5	update	update*0.022, version*0.006
6	gameplay	mode*0.004, multiplayer*0.007
7	community	community*0.065, people*0.049
8	server	server*0.025, account*0.017
9	region	russian*0.039, language*0.013
10	price	money*0.016, price*0.005
11	matchmaking	match*0.010, matchmaking*0.012
12	cheating	hacker*0.121, cheater*0.043
13	skill	bomb*0.007, nuke*0.006
14	learning curve	learning*0.005, noob*0.008
15	teamwork	team*0.019, mate*0.003
16	ranking	rank*0.007, silver*0.007

(d) PUBGm

No.	**Inferred topics**	**Keyword examples (PUBGm)**
1	graphics	graphic * 0.049, fps * 0.004
2	character	skin * 0.010, outfit* 0.007
3	map	map * 0.051, world * 0.020
4	optimization	bug * 0.031, glitch * 0.038
5	update	update * 0.096, season * 0.024
6	gameplay	gameplay * 0.005, mode*0.006
7	community	friend * 0.024, people * 0.014
8	server	server * 0.033, ping * 0.013
9	region	bangladesh * 0.003, china * 0.002
10	platform	money * 0.026, app * 0.080
11	teamwork	team * 0.012, mate *0.001
12	cheating	hacker * 0.038, cheater * 0.045
13	skill	gun * 0.008, battle* 0.013
14	learning curve	noice * 0.004, level * 0.005
15	ranking	rank * 0.008, ban * 0.018
16	device	phone * 0.034, ram * 0.016

topic reflects the spatial design and environmental storytelling that contribute to the strategic depth and replayability of a game.

The “optimization” topic is often a focal point of technical critique, addressing the efficiency of game performance across different hardware setups. The “update” topic is a testament to the game’s evolution, highlighting the community’s response to new content, balance changes, and feature additions. Lastly, the “gameplay” topic is central to the core gaming experience, encompassing mechanics, controls, and overall enjoyment of the game. Collectively, these GRTs provide a snapshot of the technical and creative efforts that go into game development and the resultant player perceptions that can influence a game’s reputation and success.

Player-related Topics (PRT) The second category, PRT, with topics numbered from 7 to 16 in Table 6.3, captures the essence of player engagement and the social aspects of gaming. These topics are reflective of the community’s pulse and the individual player’s journey within the game’s ecosystem. The “skill” topic, for example, not only encompasses the player’s proficiency but also touches upon the learning curve and the resources available for skill improvement. The “cheating” topic is indicative of the challenges faced in maintaining a fair play environment and the impact of such issues on the game’s integrity and player trust. The “server” topic’s focus on connectivity issues such as lag and server downtime speaks to the technical quality of the gaming experience, which is paramount in fast-paced, competitive play. Other PRTs address the economic aspects of gaming, such as the “price” of in-game items, which reflects the monetization strategies of game operators and their reception by the player base.

The PRTs, while offering a granular view of player engagement, present interpretative intricacies due to the specific and varied meanings attributed to each keyword across different games. These topics, despite their complexity, are instrumental in shedding light on the nuances of player behavior, strategies, preferences, and the overall gaming milieu. They serve as a barometer for player satisfaction and are indicative of areas where game developers can focus their attention to enhance the user experience. Through the lens of PRTs, we gain a deeper understanding of the esports ecosystem, including player development, community building, and the social dynamics that influence game popularity and longevity.

6.3 Topic Sentiment Analysis Results

The following subsections detail the sentiment analysis results, providing a comparative view of how different topics elicit varied emotional responses from the gaming community. The findings are instrumental in pinpointing specific areas that may require attention from developers or highlight successful features that resonate well with the audience.

6.3.1 Methodology of Sentiment Analysis

In our research, we sought to delve deeper into the sentiment analysis of esports game reviews using a sophisticated approach. We employed the "bert-base-uncased" model, renowned for its effectiveness in natural language processing tasks. This time, we enhanced the model's capability by integrating a downstream layer, specifically tailored for End-to-End Aspect-Based Sentiment Analysis (E2E-ABSA). This modification was aimed at capturing the nuanced sentiments expressed in the extensive player reviews more accurately. To accommodate the model's input size limitations, we meticulously segmented the lengthy reviews into smaller paragraphs. This ensured that each input segment remained within the model's maximum sequence length, thereby preserving the integrity of the data fed into the model for training and analysis.

Through a series of 1,500 steps, we meticulously fine-tuned the parameters, selecting the most effective model iteration based on key performance indicators: precision, recall, and the Micro-F1 score. In a bid to reinforce the validity and reliability of our model, we trained it across 30 different runs using a variety of random seeds. This strategy was instrumental in normalizing the results, providing a solid, reliable foundation for our sentiment analysis. Table 6.4 below illustrates the performances of various models tested on the esports dataset, showcasing the BERT-Transformer model's superior performance with a Micro-F1 score of 78.96%, indicating its enhanced ability to classify sentiments accurately.

This data not only evidences the BERT-Transformer's efficacy in parsing through

Table 6.4: ABSA performances on esports dataset.

Model	Precision	Recall	Micro-F1
Bi-LSTM	0.7673	0.7176	0.7416
BERT-Transformer (no fine-tuning)	0.6900	0.7587	0.7227
BERT-Liner	0.7441	0.7970	0.7696
BERT-Transformer	0.7937	0.7855	**0.7896**

sentiment-laden feedback but also validates the methodological rigor of our approach. The BERT-Transformer has affirmed its place as the backbone of our E2E-ABSA by consistently outperforming other models in micro-F1 score, setting a new benchmark in our analytical capabilities.

6.3.2 Sentiment Analysis for All Players

Fig. 6.4 presents a detailed analysis of sentiment distribution for various topics within four major esports games: PUBG(P), Dota2(D), CS:GO(C), and PUBGm(M). The sentiments have been carefully categorized into positive, neutral, and negative responses, distinguished by the colors green, yellow, and red, respectively. On the chart, the y-axis represents the percentages, illustrating the proportion of sentiments across different discussion topics for each game.

It's noteworthy that the aggregate of sentiments—positive, neutral, and negative—tends to cluster around the 5% threshold for several topics like region, learning curve, and ranking. This pattern suggests that these particular topics might not be the main focus of discussions among players, and it also indicates that a significant portion of the player reviews could contain a high level of non-informative content. This highlights a challenge in the data: the scarcity of clear and actionable feedback from players, which is essential for guiding improvements.

The chart also brings to light specific concerns and prevalent issues for each game. Concerns within PUBG are varied, with significant emphasis on map design, optimization, server performance, and cheating. This contrasts with Dota2, where the feedback predominantly pertains to community-related issues, emphasizing the importance of social dynamics in the gaming experience. In alignment with common issues in first-person shooters, CS:GO's main criticism centers on cheating, which affects gameplay integrity. PUBGm is distinctively challenged by issues with updates

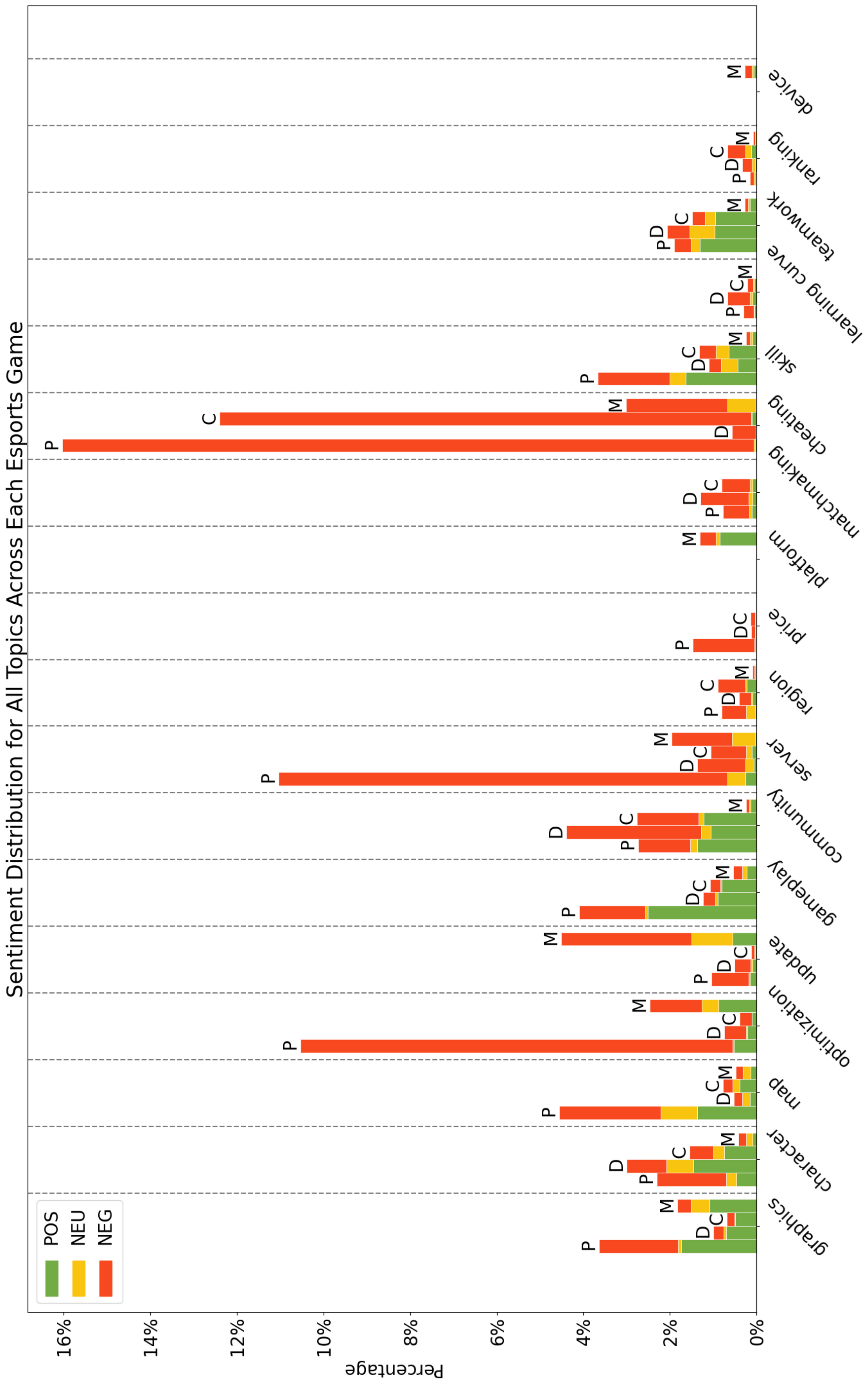


Figure 6.4: Topics sentiments distribution of all players among four games.

and cheating, underlining the unique difficulties faced in mobile gaming with regards to maintaining frequent updates and ensuring fair play.

This sentiment analysis is invaluable for developers and community managers, as it provides a concentrated overview of which areas are in need of focus and refinement. Differences in player feedback across these games suggest tailored strategies could be deployed to improve the user experience, targeting the most critical issues identified by players. For example, while common concerns like optimization and cheating require technical solutions, the focus on community issues in Dota2 indicates that player expectations extend to the social environment of the game, suggesting players have higer sensitivity towards topics of PRT group.

6.3.3 Sentiment Analysis within the Early Gaming Hours

The first hours of gameplay are a pivotal juncture that can significantly shape a player's long-term relationship with an esports game. This range represents a crucial period where players form their foundational opinions and decide whether to invest more time into a esports game. Our analysis examines the sentiment trends that emerge during these early hours on three PC esports games, providing insights into the initial challenges and draws of each game as perceived by new players. The absence of PUBGm from this novice-centric sentiment analysis is due to the lack of available playtime data, which is a crucial variable in this context.

Playtime vs. #Not_Recommended Reviews Percentage

Fig. 6.5 shows the relationship between the playtime at review and the percentage of #not_recommended reviews for three different esports games: PUBG, Dota2, and CS:GO. The x-axis is on a logarithmic scale and represents the playtime in hours at which a review was written, while the y-axis indicates the percentage of #not_recommended reviews. Here are several insights from the graph:

1. Initial Playtime (Less than 1 hour):

   - PUBG has a very high initial percentage of #not_recommended reviews when playtime is less than 1 hour. This suggests that players may be encountering

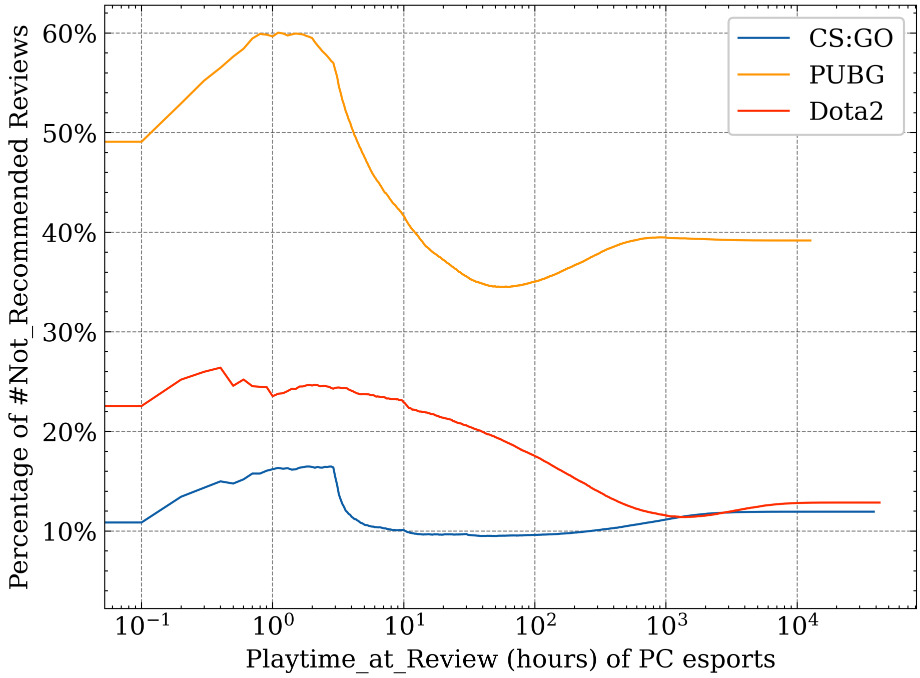


Figure 6.5: Percentage of #Not_Recommended Reviews by Playtime (hours).

immediate issues, which could be related to gameplay, performance, or other technical problems that are impacting their early game experience.

2. Early Gameplay Experience (1 to 10 hours):

 - CS:GO starts with a lower percentage of #not_recommended reviews, and this percentage decreases steadily as playtime increases, which could imply a positive early user experience that improves with more engagement.
 - Dota2 exhibits an increase in #not_recommended reviews as players spend more time up to around 10 hours, suggesting some elements of the game may not meet player expectations after they have a moderate amount of experience with it.

3. Extended Playtime Experience (10 to 100 hours and beyond):

 - The percentage of #not_recommended reviews for PUBG shows a steep decline after the initial playtime and then stabilizes after around 10 hours of play. This could indicate that players who continue beyond the initial hours find the game more enjoyable or at least less negative.
 - Dota2's #not_recommended reviews peak at around 10 hours of playtime but

then begin to decrease, suggesting that players who persist beyond this point may find aspects of the game more appealing or may adapt to its gameplay.

- For CS:GO, the trend towards fewer #not_recommended reviews continues as playtime increases, which might reflect a solid game experience that satisfies players the more they engage with the game.

4. Long-term Engagement (1,000+ hours):

- The percentage of #not_recommended reviews for all games except PUBG flattens out at longer playtimes, indicating a baseline level of dissatisfaction that persists regardless of how long players have engaged with the game. This could suggest that dedicated players have consistent grievances that are not resolved over time, or that these issues are inherent to the game experience itself.
- For PUBG, the curve flattens significantly as playtime extends beyond 100 hours, suggesting that players who invest a substantial amount of time tend to have a more consistent (and potentially better) gaming experience, or it might be that players who continue to play despite initial dissatisfaction adjust their expectations and find aspects of the game they enjoy.

Building on the initial insights, our subsequent analysis will delve deeper into the experiences of players who have engaged with the games for less than 10 hours. This examination is pivotal for unpacking the early user experience and for identifying the decisive factors that influence a player's initial commitment to a game.

Novice Player Sentiment Analysis

The analysis presented in Figure 6.6 concentrates on novice players with less than 10 hours of playtime in the popular esports titles PUBG (p), Dota2 (d), and CS:GO (c).

When examining the sentiment distribution specific to novice players, it's observed that the percentages of sentiments—positive, neutral, and negative—are significantly lower as compared to those from the general player base, with the highest percentage of sentiment expression barely reaching 8%. This data implies that novice players tend to express a more neutral to negative sentiment towards these games, with negative sentiments often being the most dominant in the clusters for each game and topic.

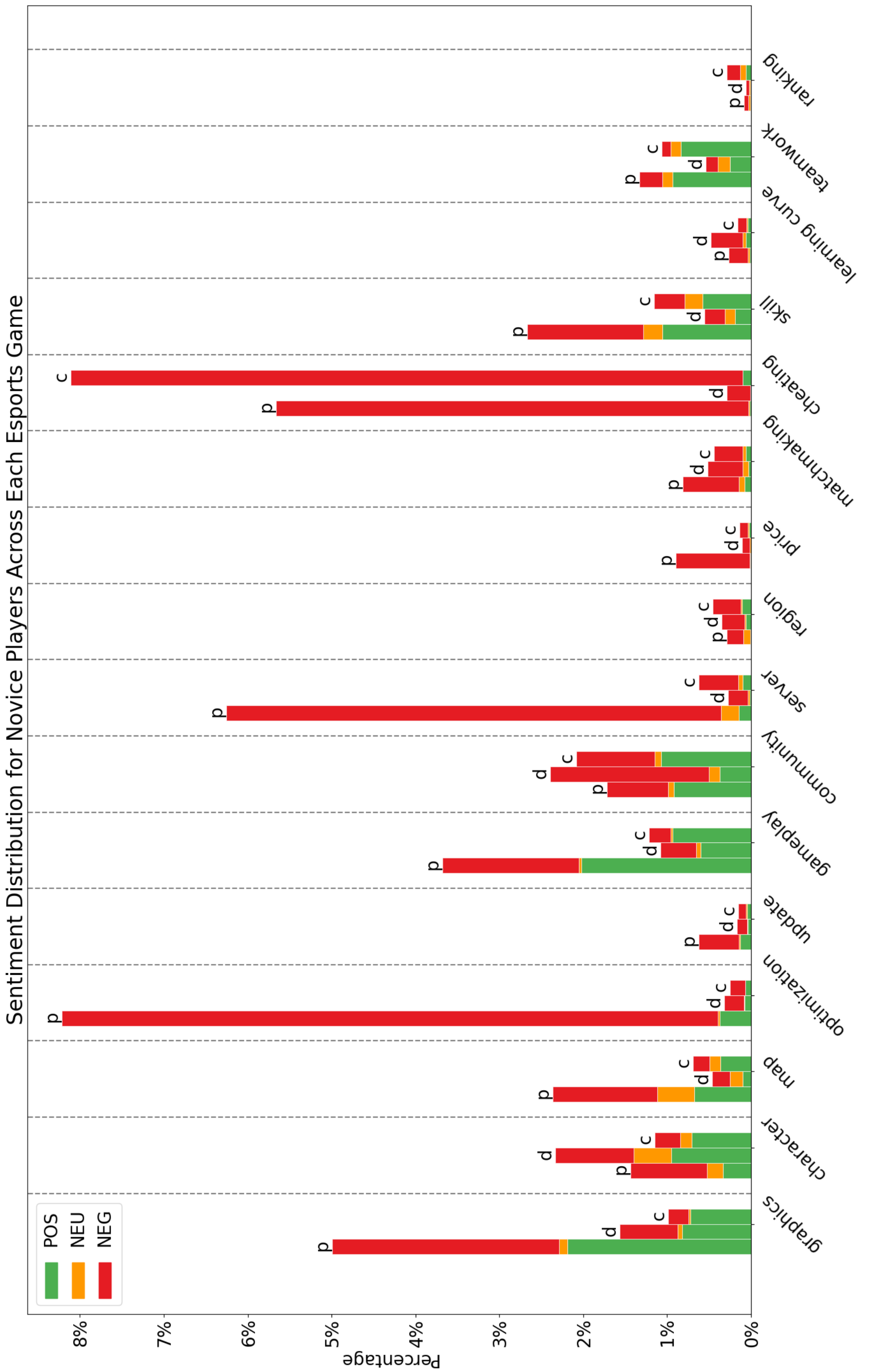


Figure 6.6: Topics sentiments distribution for early players.

Such a trend could indicate that the initial gaming experiences for these novices may be fraught with challenges or less enjoyable elements.

The comparative insights gathered from this analysis reveal a stark contrast: general sentiment intensities across all topics for all players are markedly higher than those for novice players. This disparity suggests that as players accrue more time within the game, their sentiments become more pronounced. The overall player population demonstrates a substantial increase in positive sentiments, possibly reflecting a journey where initial challenges are surmounted and gameplay proficiency is gained, leading to an enhanced gaming experience.

Drilling down into game-specific trends, the analysis identifies that novice players express significantly higher negative sentiment on certain issues. In PUBG, the graphics and gameplay have drawn considerable negative sentiment from novices, while in CS:GO, the prevalent issue of cheating is a major point of contention. These insights highlight specific areas that are particularly frustrating for new players, suggesting these are critical aspects that game developers should address to improve the initial experience for newcomers.

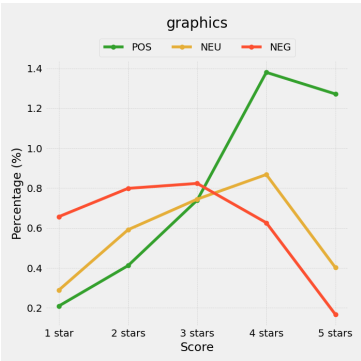
6.3.4 Sentiment Distribution Across Different Ratings

In the mobile gaming platform, similar to the approach on PC, we aim to understand the sentiment of novice players. However, without access to playtime data—an essential metric on PC—our analysis pivots to examine the trends in rating scores. This shift allows us to infer the initial impressions and experiences of players, which can be particularly telling of a game's accessibility and immediate appeal to newcomers. By studying these rating trends, we can glean insights into the first-contact sentiment, a crucial indicator of a game's ability to engage and satisfy its burgeoning player base.

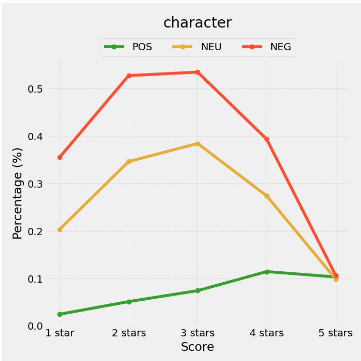
Game-related Topics (GRT)

The sentiment analysis across GRT groups in PUBGm, as captured in Fig. 6.7, reveals a multifaceted relationship between player ratings and their feedback. The data indicates that players' initial interactions with game elements such as characters, gameplay mechanics, and visual fidelity elicit diverse emotional responses that correlate with their rating behavior.

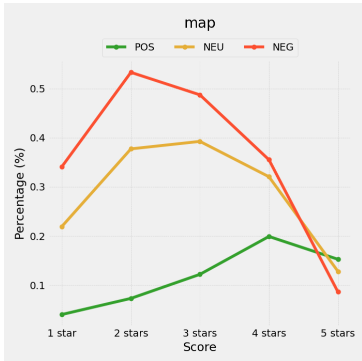
A deeper dive into these sentiment trends showcases that player satisfaction is not linearly tied to rating scores. Positive feedback peaks at specific rating thresholds, suggesting that players have nuanced expectations for each game aspect that are not always met, even at higher ratings. Furthermore, the substantial negative sentiment associated with certain features at lower ratings highlights areas where the game may fall short in delivering a satisfactory user experience to newcomers.



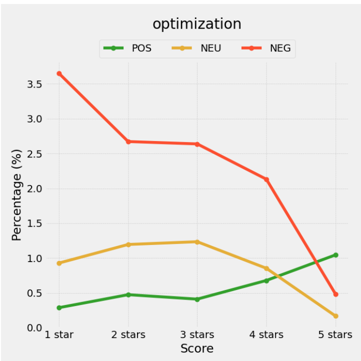
(a) graphics



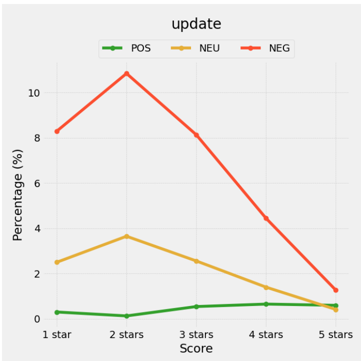
(b) character



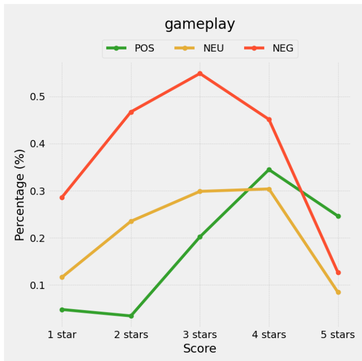
(c) map



(d) optimization



(e) update



(f) gameplay

Figure 6.7: GRT group Sentiments for PUBGm.

Sentiment analysis within game-related topics presents a complex picture of player satisfaction. Negative sentiment towards character design follows a significant bell curve across the ratings, with a peak at around 0.55% for 3-star ratings. This suggests that players have high expectations for character development that are not always fulfilled. In gameplay, the peak of positive sentiment reaches about 0.4% at 4-star ratings, yet there's a noticeable decline at 5-stars, indicating that players may not feel completely satisfied with gameplay even in higher-rated reviews. Graphics receive a substantial amount of positive sentiment, the highest being around 1.4% at 4-star ratings, but see a decline at the top rating, which may point to a gap between players' expectations and the actual visual experience delivered by the games.

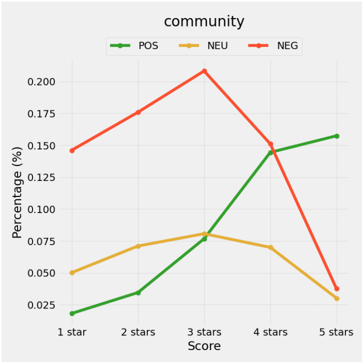
When it comes to map, discussions maintain a moderate sentiment distribution but with the negative sentiment all over 0.34% before 5-stars, hinting at wide player criticism. Optimization issues are particularly glaring, with negative sentiment reaching over 3.5% at the 1-star rating, indicating that technical performance is a significant driver of player dissatisfaction. Similarly, updates trigger a notable spike in negative sentiment at around 10% for 2-star ratings, underscoring the importance of how game changes and improvements are managed and perceived.

These detailed sentiments illustrate that while players may generally rate games positively, there's a discernible undercurrent of specific critiques. This complexity in player satisfaction is characterized by the fact that exceptional aspects of a game like character and graphics are often celebrated at 4-stars but do not necessarily translate into perfection at 5-stars. At the same time, pivotal game aspects such as optimization and updates can significantly sway player sentiment, becoming key areas for developers to focus on for enhancement.

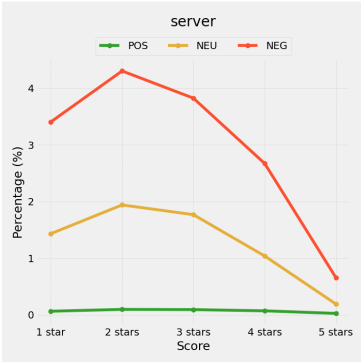
Player-related Topics (PRT)

The comprehensive feedback collected from the PUBGm community is a treasure trove of insights, offering a window into the game's strengths and potential areas for improvement. An in-depth sentiment distribution analysis, as illustrated in Fig. 6.8, sheds light on the various nuances of player experiences within the Player-related Topics (PRT) group. Key areas of focus—ranging from server performance to device compatibility—have been highlighted by players, providing a roadmap for targeted enhancements.

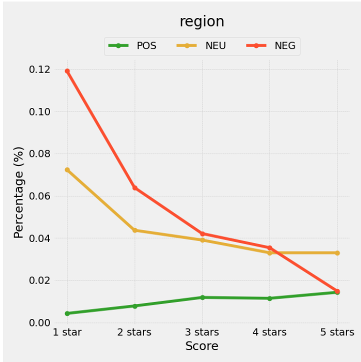
In the landscape of player sentiment within PUBGm, certain topics stand out for their pronounced impact on player satisfaction. The issue of cheating is paramount, with a striking 7% of negative sentiment in 1-star ratings, overshadowing other concerns and indicating a critical area where player expectations clash with their experiences. Similarly, server-related issues command attention with nearly 4% negative sentiment at the similar rating level, underscoring the importance of technical stabil-



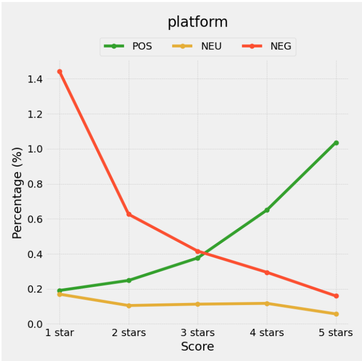
(a) community



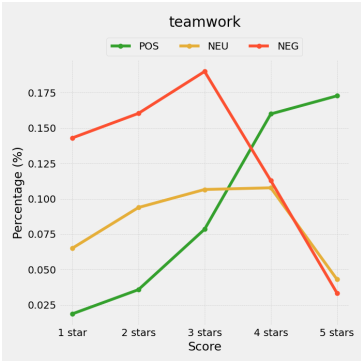
(b) server



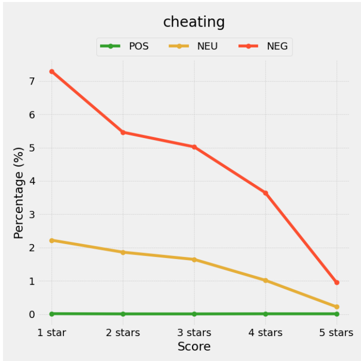
(c) region



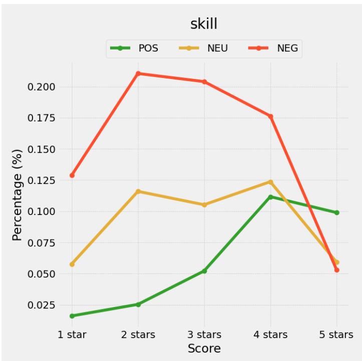
(d) platform



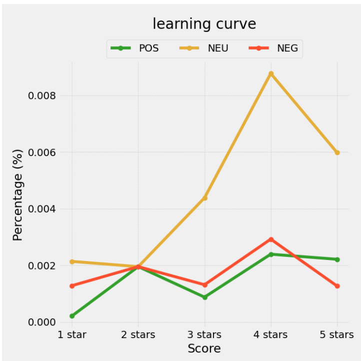
(e) teamwork



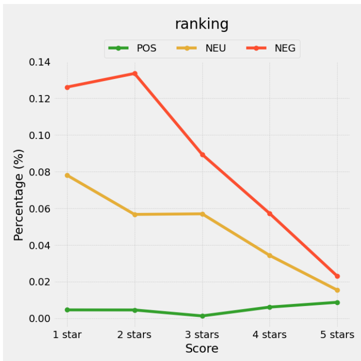
(f) cheating



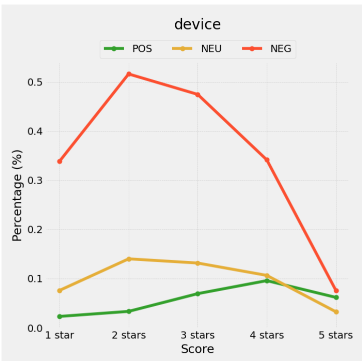
(g) skill



(h) learning curve



(i) ranking



(j) device

Figure 6.8: PRT group Sentiments for PUBGm.

ity in player retention and satisfaction.

Device performance, while eliciting a significant response with close to 0.5% negative sentiment in 2-stars reviews, suggests that technical compatibility is a notable barrier for some players, yet it does not ignite as widespread a concern as cheating or server issues. Community engagement, peaking at around 0.2% in negative sentiment, indicates a nuanced area of player experience that, while essential, is often overshadowed by more immediate gameplay-related issues.

Topics like learning curve and teamwork register lower percentages of negative sentiment, suggesting these aspects, although integral to the game's social and competitive framework, may influence player perception more subtly. Positive sentiment in these areas peaks at intermediate ratings, reflecting players' recognition of their importance to the overall game experience, yet without the intensity of response that more critical issues provoke.

Skill balancing and matchmaking are also pivotal, as indicated by the peak in negative sentiment at higher ratings. This suggests a critical discourse among players who are invested in the game but feel that the skill systems may not effectively cater to a fair and balanced competitive environment. Meanwhile, regional issues, while less prominent overall, still elicit a considerable response at lower ratings, pointing to the importance of accessibility and equitable play conditions across different geographic locations.

In a broader context, while all player-related topics contribute to the fabric of player sentiment, those with higher percentages like cheating and server performance represent urgent areas for developer intervention. In contrast, while topics such as device compatibility and community sentiment register a smaller volume of feedback, they remain integral to crafting a comprehensive and engaging player experience. Addressing these concerns holistically can lead to improvements that resonate across the entire spectrum of player engagement, from the novice encountering the game for the first time to the seasoned player deeply invested in the game's ecosystem.

Chapter 7

Discussion

This chapter expands upon the methodological and data availability contributions detailed in the dissertation, highlighting the broader applicability and potential adaptations of these methodologies across various domains.

7.1 General Applicability and Adaptation

The methodological advancements developed within this dissertation, emphasizing specialized topic modeling and sentiment analysis, are poised for impactful applications across a spectrum of industries reliant on user-generated content. The study's core focus on esports game reviews serves as a foundational case, demonstrating the effectiveness and adaptability of these methodologies. However, the intrinsic value of these analytical frameworks extends far beyond this niche, providing a versatile toolset for extracting actionable insights from diverse textual data sources.

The depth and adaptability of these methodologies make them particularly well-suited for various sectors, each with unique challenges and opportunities for leveraging user-generated content. Below, we delve into specific applications across different domains, illustrating the broad utility and transformative potential of the dissertation's methodological contributions:

E-Commerce Applications: The burgeoning e-commerce sector, where user reviews can significantly influence purchasing decisions, stands to gain considerable benefits from enhanced topic modeling and sentiment analysis. By dissecting customer feedback at scale, businesses can uncover granular insights into consumer sentiment, preferences, and expectations, enabling more nuanced market understanding and customer engagement strategies.

Educational Platforms: Similarly, in the educational domain, the ability to systematically analyze student feedback can provide educators and administrators with critical insights into course effectiveness, student engagement, and overall satisfaction. Such analysis can inform targeted improvements, enhancing the educational experience and outcomes.

Healthcare Feedback Analysis: In healthcare, analyzing patient feedback through advanced topic modeling and sentiment analysis can illuminate patient experiences, expectations, and areas for service improvement. These insights can drive patient-centered care initiatives, enhance service quality, and inform healthcare policies and practices.

Social Media Monitoring: The pervasive influence of social media platforms in shaping public opinion and trends underscores the need for sophisticated content analysis tools. The methodologies developed in this dissertation can be applied to monitor, analyze, and interpret social media discourse, providing valuable insights for content strategy, crisis management, and community engagement.

7.2 Technical and Social Significance

This research contributes significantly to both the technical and social dimensions of data analysis, particularly in understanding user-generated content. It bridges the gap between complex data mining techniques and practical, user-centered applications, highlighting the dual impact of technical advancements on society.

- **Technical Significance:** The technical merit of this dissertation lies in its innovative application of natural language processing (NLP) techniques to esports game reviews, a domain characterized by rich but unstructured data. By developing a specialized model that combines topic modeling with sentiment analysis, this research tackles the challenge of extracting meaningful insights from vast amounts of text. This methodological advancement not only enhances the precision of sentiment analysis but also broadens the scope of topic modeling, making it a valuable tool for analyzing diverse datasets beyond esports, such as customer feedback in retail or patient comments in healthcare. Through its contribution to the fields of data mining and machine learning, the dissertation not only addresses specific challenges within esports but also sets a precedent for future research in applying these techniques to understand and leverage unstructured data across various sectors.

- **Social Significance:** Beyond the technical contributions, the social implications of this research are profound. By enhancing our ability to interpret and act on user sentiment, the methodologies developed in this dissertation have the potential to transform user experience across multiple platforms. This transformation is achieved by fostering environments that are not only responsive to user feedback but also anticipatory of user needs. In esports, this means creating gaming experiences that are more engaging, inclusive, and satisfying for players. More broadly, the application of these methodologies can lead to more user-centric services and products, from social media platforms that better cater to user well-being to e-commerce sites that can more accurately predict and meet consumer desires. This shift towards a more responsive and user-informed approach has the potential to redefine the relationship between service providers and users, making it more dynamic, symbiotic, and driven by mutual value creation.

- **Broadening the Impact:** The transition of esports from a niche interest to a mainstream entertainment phenomenon illustrates the transformative power of advanced analytics. This transition is reflective of a larger movement towards data-driven decision-making across industries. By demonstrating how sophisticated analysis of user-generated content can yield actionable insights, this research contributes to a broader understanding of how data can inform strategy, innovation, and user engagement. In a world increasingly dominated by digital interactions, the ability to derive meaningful insights from user data is not just a competitive advantage; it's a necessity for sustainability and growth.

Chapter 8

Conclusion & Future Work

8.1 Summary of Findings

This comprehensive analysis delves into player experiences and preferences across both PC and mobile esports platforms, with a focus on games like PUBG, Dota2, and CS:GO, along with the mobile version PUBGm.

In the realm of PC gaming, significant trends emerge. Key insights include the high value players place on graphics, gameplay, and character design, as evident in Game-related Topics (GRT). Player-related Topics (PRT) vary by game, with issues like community in Dota2 and cheating in PUBG and CS:GO. The sentiment analysis suggests a shift in esports game operators' approach from product-centric to service-centric, particularly addressing negative sentiments more prevalent in PRT topics. The feedback from novice players underlines the importance of enhancing the gaming experience within the first 10 hours of gameplay to boost engagement and player retention.

Transitioning to mobile esports, PUBGm stands as a testament to the capabilities of mobile gaming but also mirrors challenges found in its PC counterpart. The analysis of 6 million English reviews for PUBGm indicates similarities in issues, with an added emphasis on device compatibility. While graphical excellence in PUBGm is widely acclaimed, diverse opinions on character and map designs reflect the challenge of catering to a global audience. The need for technical optimization, updates, and robust anti-cheat systems are as crucial in mobile as in PC gaming. For PUBGm, server performance and compatibility across a range of mobile devices emerge as significant concerns.

The comprehensive analysis across both PC and mobile gaming platforms yields several pivotal insights into the esports gaming sector:

1. Reviews of esports games are an invaluable trove for discovering prevalent issues. Our analytical results reveal that, regardless of whether on PC or mobile platforms, negative commentary on elements of esports games generally overshadows other sentiments. This prevalence of critical feedback in esports reviews is instrumental for uncovering areas of player dissatisfaction and concern. Among the issues, aspects such as game optimization, community dynamics, server reliability, and cheating concerns are universally prominent, consistently capturing the attention of the player communities.

2. Delving into data from novice players on the PC platform, we find that these players often exhibit a spectrum of sentiment that skews from neutral to negative. Notably, negative sentiments tend to dominate the discourse for each game and topic cluster. Beyond this, novice players articulate a greater number of issues, including those about graphics and gameplay—areas that do not typically stand out among the broader player demographics. This implies a distinctive set of challenges faced by new entrants, potentially affecting their initial engagement and long-term retention.

3. On the mobile platform, the absence of playtime data necessitated a pivot towards analyzing rating trends to explore differences in feedback among diverse player segments. Beyond the general player concerns identified—such as optimization, community interaction, server functionality, and instances of cheating—players of PUBGm on mobile platforms expressed pronounced discontent with game updates. This finding is particularly salient, signifying an area that warrants closer examination and responsive action from game developers.

In juxtaposition with prior research on existing problems in the esports environment, our quantitatively enriched analysis provides a deeper understanding of this domain:

- The learning curve is acknowledged as an obstacle, yet it does not command a significant share of the conversation. Even for a game with a steep learning curve like Dota2, comments explicitly addressing this issue constitute less than 1% of the total commentary.

- The experimental results affirm the reality of a less-than-welcoming environment for novice players, with their feedback revealing a notably higher proportion of negative sentiments compared to the collective player base.3

- The present analysis points out that Dota2 players place a heightened emphasis on community aspects compared to other esports titles. We infer this to be a reflection of Dota2's intensely team-centric gameplay, where effective collaboration and teamwork are crucial for successful play.

- For game updates, significant emphasis has been placed on this aspect in the context of the mobile platform, particularly with PUBGm. In contrast, our analysis of PC esports titles does not indicate an equivalent level of concern, suggesting platform-specific variances in player priorities and the perceived impact of updates.

8.2 Conclusion

This dissertation unveils an innovative data analysis framework tailored specifically for esports game reviews, addressing the intricate challenge posed by the vast and diverse array of player feedback. This framework's concentrated focus on the nuances of esports content signifies a critical advancement toward a more sophisticated and effective analysis of player feedback.

In this research, we meticulously gathered a comprehensive dataset encompassing approximately eight million reviews from key esports games, drawn from both the Steam and Google Play platforms. This extensive collection, enriched by its diverse platform origins, underpins the robustness and broad applicability of our experimental outcomes. Our proposed framework integrating enhanced topic modeling with sentiment analysis, significantly elevates the accuracy in pinpointing player emotions and dissecting various dimensions of game dynamics. Through comparative analysis with existing models, our framework's superior efficacy is vividly demonstrated, reinforcing the essential role of each of its components in crafting a holistic analytical approach.

The insights gleaned from this study are pivotal in decoding the complex tapestry of esports players' feedback, especially in spotlighting recurrent issues that resonate across player communities. The methodologies and tools developed in this research hold immense potential for game developers and community managers. They offer refined, in-depth perspectives on player feedback, which can serve as a cornerstone for shaping strategies around game updates and community engagement initiatives. These insights are particularly crucial in an era where player-centric approaches are paramount in game development and community building.

While the model showcased in this dissertation harbors substantial promise and stands as a testament to the potential of data-driven approaches in esports analytics, it also marks just the beginning of a broader journey. It lays a solid foundation that beckons future academic inquiries to not only refine and augment its capabilities but also to innovate further in alignment with the ever-evolving dynamics of the esports industry. As the sector continues to grow and player feedback becomes increasingly pivotal in game development processes, the need for advanced, nuanced analytical tools such as the one presented in this study becomes ever more pressing. This research, therefore, not only contributes to the current academic discourse but also paves the way for future explorations, aiming to continually enhance the symbiotic relationship between game developers and their vibrant player communities.

8.3 Limitations & Future Work

This research provides an in-depth analysis of player experiences in PC and mobile esports, acknowledging certain limitations that are pivotal for future work. The focus on English-language reviews in this study restricts the cultural and linguistic diversity of the analyzed player feedback, thus excluding significant non-English speaking markets and potentially missing vital insights from regions with a strong esports presence. The confinement to specific platforms like Steam for PC games and Android for mobile games also limits perspectives from other critical platforms such as Metacritic and the iOS ecosystem. Additionally, the geographical coverage falls short in capturing the distinctive gaming cultures in crucial areas like China and Southeast Asia. The selection of games, while covering popular titles, omitted other major esports games like League of Legends, limiting the variety of gaming genres and styles examined.

Future research aims to expand inclusivity and coverage in esports gaming analysis. To address the linguistic limitations and capture a more comprehensive reflection of the global esports community, plans include broadening the research scope to encompass a variety of languages and a wider array of popular games. Incorporating data from diverse platforms, including iOS and sites like Metacritic, will offer a more complete view of player sentiments across different user bases. A deeper dive into the unique gaming cultures of regions like China and Southeast Asia is essential due to their distinct player dynamics and significant growth in the esports sector. Exploring mobile MOBAs, a substantial part of mobile esports, and a thorough cross-platform analysis are also planned. These efforts are directed toward enriching the understanding of the mobile esports landscape and investigating specific game cultures and strategic hotspots in the gaming world. The anticipated findings from this expanded research approach are expected to significantly enhance the services and strategies of esports operators, allowing them to tailor their offerings more effectively to the diverse and evolving needs of the global gaming community.

The methodological approach developed in this study, while primarily applied to English-language reviews, holds general applicability across languages and cultures. However, specific adaptations may be necessary to accommodate linguistic and cultural nuances in non-English-speaking markets. The data collection process will need to identify and access user communities in various languages, considering the availability of public APIs and platforms. The preprocessing of data, a common step in NLP and data analysis, may not require redevelopment but will need adjustments for different languages. The core methodologies and techniques proposed in this dissertation, including aspect-based sentiment analysis, appear broadly applicable across languages. Still, future work should explore their effectiveness in diverse linguistic contexts, aiming to maintain as much generality as possible in the approach while ensuring accuracy and cultural relevance in the analysis.

Bibliography

[1] Newzoo, "Global esports market report. light version," 2019. Retrieved from https://newzoo.com/resources/trend-reports/newzoo-global-esports-market-report-2019-light-version.

[2] Newzoo, "Global games market report," 2020. Retrieved from https://newzoo.com/resources/trend-reports/newzoo-global-games-market-report-2020-light-version.

[3] Newzoo, "Global esports live streaming market report," 2022. Retrieved from https://newzoo.com/resources/trend-reports/newzoo-global-esports-live-streaming-market-report-2022-free-version.

[4] G. Cheung and J. Huang, "Starcraft from the stands: understanding the game spectator," in *Proceedings of the SIGCHI conference on human factors in computing systems*, pp. 763–772, 2011. DOI:10.1145/1978942.1979053.

[5] F. Bányai, M. D. Griffiths, O. Király, and Z. Demetrovics, "The psychology of esports: A systematic literature review," *Journal of gambling studies*, vol. 35, pp. 351–365, 2019. DOI:10.1007/s10899-018-9763-1.

[6] T. L. Taylor, *Watch Me Play: Twitch and the Rise of Game Live Streaming*. Princeton: Princeton University Press, 2019. DOI:10.1515/9780691184975.

[7] Y. Kou and X. Gui, "Playing with strangers: understanding temporary teams in league of legends," in *Proceedings of the first ACM SIGCHI annual symposium on Computer-human interaction in play*, pp. 161–169, 2014. DOI:10.1145/2658537.2658538.

[8] J. Macey and J. Hamari, "Investigating relationships between video gaming,

spectating esports, and gambling," *Computers in Human Behavior*, vol. 80, pp. 344–353, 2018. DOI:10.1016/j.chb.2017.11.027.

[9] Y. Seo and S.-U. Jung, "Beyond solitary play in computer games: The social practices of esports," *Journal of Consumer Culture*, vol. 16, no. 3, pp. 635–655, 2016. DOI:10.1177/1469540514553711.

[10] W. A. Hamilton, O. Garretson, and A. Kerne, "Streaming on twitch: fostering participatory communities of play within live mixed media," in *Proceedings of the SIGCHI conference on human factors in computing systems*, pp. 1315–1324, 2014. DOI:10.1145/2556288.2557048.

[11] M. Pontiki, D. Galanis, H. Papageorgiou, I. Androutsopoulos, S. Manandhar, M. Al-Smadi, M. Al-Ayyoub, Y. Zhao, B. Qin, O. De Clercq, *et al.*, "Semeval-2016 task 5: Aspect based sentiment analysis," in *10th International Workshop on Semantic Evaluation (SemEval 2016)*, 2016. DOI:10.18653/v1/S16-1002.

[12] K. Yauris and M. L. Khodra, "Aspect-based summarization for game review using double propagation," in *2017 International Conference on Advanced Informatics, Concepts, Theory, and Applications (ICAICTA)*, pp. 1–6, IEEE, 2017. DOI:10.1109/ICAICTA.2017.8090997.

[13] B. Strååt and H. Verhagen, "Using user created game reviews for sentiment analysis: A method for researching user attitudes," in *GHITALY@ CHItaly*, 2017.

[14] B. Strååt, H. Verhagen, and H. Warpefelt, "Probing user opinions in an indirect way: an aspect based sentiment analysis of game reviews," in *Proceedings of the 21st International Academic Mindtrek Conference*, pp. 1–7, 2017. DOI:10.1145/3131085.3131121.

[15] N. Alturaief, H. Aljamaan, and M. Baslyman, "Aware: Aspect-based sentiment analysis dataset of apps reviews for requirements elicitation," in *2021 36th IEEE/ACM International Conference on Automated Software Engineering Workshops (ASEW)*, pp. 211–218, IEEE, 2021. DOI:10.1109/ASEW52652.2021.00049.

[16] M. K. Baowaly, Y.-P. Tu, and K.-T. Chen, "Predicting the helpfulness of game reviews: A case study on the steam store," *Journal of Intelligent & Fuzzy Systems*, vol. 36, no. 5, pp. 4731–4742, 2019. DOI:10.3233/JIFS-179022.

[17] S. Ruseti, M.-D. Sirbu, M. A. Calin, M. Dascalu, S. Trausan-Matu, and G. Militaru, "Comprehensive exploration of game reviews extraction and opinion mining using nlp techniques," in *Fourth International Congress on Information and Communication Technology*, (Singapore), pp. 323–331, Springer Singapore, 2020. DOI:10.1007/978-981-15-0637-6_27.

[18] M. Hu, Y. Peng, Z. Huang, D. Li, and Y. Lv, "Open-domain targeted sentiment analysis via span-based extraction and classification," in *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pp. 537–546, Association for Computational Linguistics, 2019. DOI:10.18653/v1/P19-1051.

[19] X. Ma and E. Hovy, "End-to-end sequence labeling via bi-directional lstm-cnns-crf," in *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, (Berlin, Germany), pp. 1064–1074, Association for Computational Linguistics, 2016. DOI:10.18653/v1/P16-1101.

[20] M. Schmitt, S. Steinheber, K. Schreiber, and B. Roth, "Joint aspect and polarity classification for aspect-based sentiment analysis with end-to-end neural networks," in *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pp. 1109–1114, 2018. DOI:10.18653/v1/D18-1139.

[21] Y. Zhang, J. Du, X. Ma, H. Wen, and G. Fortino, "Aspect-based sentiment analysis for user reviews," *Cognitive Computation*, vol. 13, no. 5, pp. 1114–1127, 2021. DOI:10.1007/s12559-021-09855-4.

[22] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pp. 4171–4186, Association for Computational Linguistics, 2019. DOI:10.18653/v1/N19-1423.

[23] I. M. Urriza and M. A. A. Clariño, "Aspect-based sentiment analysis of user created game reviews," in *2021 24th Conference of the Oriental COCOSDA International Committee for the Co-ordination and Standardisation of Speech Databases and Assessment Techniques (O-COCOSDA)*, pp. 76–81, IEEE, 2021. DOI:10.1109/O-COCOSDA202152914.2021.9660559.

[24] S. Gunathilaka and N. De Silva, "Aspect-based sentiment analysis on mobile application reviews," in *2022 22nd International Conference on Advances in ICT for Emerging Regions (ICTer)*, pp. 183–188, IEEE, 2022. DOI:10.1109/ICTer58063.2022.10024070.

[25] X. Li, L. Bing, W. Zhang, and W. Lam, "Exploiting BERT for end-to-end aspect-based sentiment analysis," in *Proceedings of the 5th Workshop on Noisy User-generated Text (W-NUT 2019)*, pp. 34–41, 2019. DOI:10.18653/v1/D19-5505.

[26] A. Agarwal and J. Sabharwal, "End-to-end sentiment analysis of twitter data," in *Proceedings of the Workshop on Information Extraction and Entity Analytics on Social Media Data*, pp. 39–44, 2012.

[27] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," *Journal of machine Learning research*, vol. 3, no. Jan, pp. 993–1022, 2003. DOI:10.5555/944919.944937.

[28] J. Huang, S. Rogers, and E. Joo, "Improving restaurants by extracting subtopics from yelp reviews," *iConference 2014 (Social Media Expo)*, 2014. DOI:https://hdl.handle.net/2142/48832.

[29] H. Samy, A. Helmy, and N. Ramadan, "Aspect-based sentiment analysis of mobile apps reviews using class association rules and lda," in *2021 Tenth International Conference on Intelligent Computing and Information Systems (ICICIS)*, pp. 183–189, IEEE, 2021. DOI:10.1109/ICICIS52592.2021.9694242.

[30] J. Joung and H. M. Kim, "Automated keyword filtering in latent dirichlet allocation for identifying product attributes from online reviews," *Journal of Mechanical Design*, vol. 143, no. 8, p. 084501, 2021. DOI:10.1115/1.4048960.

[31] Y. Guo, F. Wang, C. Xing, and X. Lu, "Mining multi-brand characteristics from online reviews for competitive analysis: A brand joint model using latent dirichlet allocation," *Electronic Commerce Research and Applications*, vol. 53, p. 101141, 2022. DOI:10.1016/j.elerap.2022.101141.

[32] S. Feuerriegel, A. Ratku, and D. Neumann, "Analysis of how underlying topics in financial news affect stock prices using latent dirichlet allocation," in *2016 49th Hawaii International Conference on System Sciences (HICSS)*, pp. 1072–1081, IEEE, 2016. DOI:10.1109/HICSS.2016.137.

[33] H. Edison and H. Carcel, "Text data analysis using latent dirichlet allocation: An application to fomc transcripts," *Applied Economics Letters*, vol. 28, no. 1, pp. 38–42, 2021. DOI:10.1080/13504851.2020.1730748.

[34] T. Santos, F. Lemmerich, M. Strohmaier, and D. Helic, "What's in a review: Discrepancies between expert and amateur reviews of video games on metacritic," *Proceedings of the ACM on human-computer interaction*, vol. 3, no. CSCW, pp. 1–22, 2019. DOI:10.1145/3359242.

[35] X. Wang and D. H.-L. Goh, "Components of game experience: An automatic text analysis of online reviews," *Entertainment Computing*, vol. 33, p. 100338, 2020. DOI:10.1016/j.entcom.2019.100338.

[36] Y. Heng, Z. Gao, Y. Jiang, and X. Chen, "Exploring hidden factors behind online food shopping from amazon reviews: A topic mining approach," *Journal of Retailing and Consumer Services*, vol. 42, pp. 161–168, 2018. DOI:10.1016/j.jretconser.2018.02.006.

[37] A. Poushneh and R. Rajabi, "Can reviews predict reviewers' numerical ratings? the underlying mechanisms of customers' decisions to rate products using latent dirichlet allocation (lda)," *Journal of Consumer Marketing*, vol. 39, no. 2, pp. 230–241, 2022. DOI:10.1108/JCM-09-2020-4114.

[38] I. Putri and R. Kusumaningrum, "Latent dirichlet allocation (lda) for sentiment analysis toward tourism review in indonesia," in *Journal of Physics: Conference Series*, vol. 801, p. 012073, IOP Publishing, 2017. DOI:10.1088/1742-6596/801/1/012073.

[39] T. Tran, H. Ba, and V.-N. Huynh, "Measuring hotel review sentiment: An aspect-based sentiment analysis approach," in *Integrated Uncertainty in Knowledge Modelling and Decision Making*, (Cham), pp. 393–405, Springer International Publishing, 2019. DOI:10.1007/978-3-030-14815-7_33.

[40] I. Sutherland, Y. Sim, S. K. Lee, J. Byun, and K. Kiatkawsin, "Topic modeling of online accommodation reviews via latent dirichlet allocation," *Sustainability*, vol. 12, no. 5, p. 1821, 2020. DOI:10.3390/su12051821.

[41] A. Afaq, L. Gaur, and G. Singh, "A latent dirichlet allocation technique for opinion mining of online reviews of global chain hotels," in *2022 3rd International Conference on Intelligent Engineering and Management (ICIEM)*, pp. 201–206, IEEE, 2022. DOI:10.1109/ICIEM54221.2022.9853114.

[42] B. Gifford, *Reviewing the critics: Examining popular video game reviews through a comparative content analysis*. PhD thesis, School of Communication, Cleveland State University, 2013. DOI:10.18653/v1/D19-5505.

[43] I. Busurkina, V. Karpenko, E. Tulubenskaya, and D. Bulygin, "Game experience evaluation. a study of game reviews on the steam platform," in *International Conference on Digital Transformation and Global Society*, pp. 117–127, Springer, 2020. DOI:10.1007/978-3-030-65218-0_9.

[44] M. Pielka, R. Sifa, R. Ramamurthy, C. Ojeda, and C. Bauckhage, "A community detection based approach for exploring patterns in player reviews," in *Intelligent Systems and Applications: Proceedings of the 2019 Intelligent Systems Conference (IntelliSys) Volume 1*, pp. 553–565, Springer, 2020. DOI:10.1007/978-3-030-29516-5_43.

[45] D. Lin, C.-P. Bezemer, Y. Zou, and A. E. Hassan, "An empirical study of game reviews on the steam platform," *Empirical Software Engineering*, vol. 24, no. 1, pp. 170–207, 2019. DOI:10.1007/s10664-018-9627-4.

[46] J. P. Zagal, N. Tomuro, and A. Shepitsen, "Natural language processing in game studies research: An overview," *Simulation & Gaming*, vol. 43, no. 3, pp. 356–373, 2012. DOI:10.1177/104687811142256.

[47] M. Bond and R. Beale, "What makes a good game? using reviews to inform design," *People and Computers XXIII Celebrating People and Technology*, pp. 418–422, 2009. DOI:10.5555/1671011.1671065.

[48] I. J. Livingston, L. E. Nacke, and R. L. Mandryk, "The impact of negative game reviews and user comments on player experience," in *Proceedings of the 2011 ACM SIGGRAPH Symposium on Video Games*, pp. 25–29, 2011. DOI:10.1145/2037692.2037697.

[49] X. Li, Z. Zhang, and K. Stefanidis, "A data-driven approach for video game playability analysis based on players' reviews," *Information*, vol. 12, no. 3, p. 129, 2021. DOI:10.3390/info12030129.

[50] L. Ghosh, S. Saha, and A. Konar, "Bi-directional long short-term memory model to analyze psychological effects on gamers," *Applied Soft Computing*, vol. 95, p. 106573, 2020.

[51] M. Sivakumar and S. R. Uyyala, "Aspect-based sentiment analysis of mobile phone reviews using lstm and fuzzy logic," *International Journal of Data Science and Analytics*, vol. 12, no. 4, pp. 355–367, 2021.

[52] D. Hermans, "How can we create legitimate positive reviews for video games using unsupervised processing methods and ai?," 2022.

[53] J. Song, "Aspect-based sentiment analysis on mobile game reviews using deep learning," 2020.

[54] J. Newman, *Videogames*. Media Studies-New Media, Routledge, 2013.

[55] T. L. Taylor, *Raising the stakes: E-sports and the professionalization of computer gaming*. Mit Press, 2012. DOI:10.7551/mitpress/8624.001.0001.

[56] E. Witkowski, "On the digital playing field: How we "do sport" with networked computer games," *Games and Culture*, vol. 7, no. 5, pp. 349–374, 2012. DOI:10.1177/1555412012454222.

[57] Y. Seo, "Electronic sports: A new marketing landscape of the experience economy," *Journal of marketing management*, vol. 29, no. 13-14, pp. 1542–1560, 2013. DOI:10.1080/0267257X.2013.822906.

[58] J. Cohen, "A coefficient of agreement for nominal scales," *Educational and psychological measurement*, vol. 20, no. 1, pp. 37–46, 1960. DOI:10.1177/001316446002000104.

[59] D. Mimno, H. Wallach, E. Talley, M. Leenders, and A. McCallum, "Optimizing semantic coherence in topic models," in *Proceedings of the 2011 conference on empirical methods in natural language processing*, pp. 262–272, 2011. DOI:10.5555/2145432.2145462.

[60] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997. DOI:10.1162/neco.1997.9.8.1735.

[61] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," *Advances in neural information processing systems*, vol. 30, 2017. DOI:10.48550/arXiv.1706.03762.

[62] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016. DOI:10.1109/CVPR.2016.90.

[63] D. Newman, J. H. Lau, K. Grieser, and T. Baldwin, "Automatic evaluation of topic coherence," in *Human language technologies: The 2010 annual conference of the North American chapter of the association for computational linguistics*, pp. 100–108, 2010. DOI:10.5555/1857999.1858011.

[64] J. Hamari and M. Sjöblom, "What is esports and why do people watch it?," *Internet research*, vol. 27, no. 2, pp. 211–232, 2017. DOI:10.2139/ssrn.2686182.

[65] M. Vesa, J. Hamari, J. T. Harviainen, and H. Warmelink, "Computer games and organization studies," *Organization Studies*, vol. 38, no. 2, pp. 273–284, 2017. DOI:10.1177/017084061666324.

[66] I. Von Hilvoorde and N. Pot, "Embodiment and fundamental motor skills in esports, sport," *Ethics and Philosophy*, vol. 10, no. 1, pp. 14–27, 2016. DOI:10.1080/17511321.2016.1159246.

[67] F. Bányai, M. D. Griffiths, Z. Demetrovics, and O. Király, "The mediating effect of motivations between psychiatric distress and gaming disorder among esport

gamers and recreational gamers," *Comprehensive psychiatry*, vol. 94, p. 152117, 2019. DOI:10.1016/j.comppsych.2019.152117.

[68] J. G. Reitman, M. J. Anderson-Coto, M. Wu, J. S. Lee, and C. Steinkuehler, "Esports research: A literature review," *Games and Culture*, vol. 15, no. 1, pp. 32–50, 2020. DOI:10.1177/1555412019840892.

[69] G. Freeman and D. Y. Wohn, "Social support in esports: building emotional and esteem support from instrumental support interactions in a highly competitive environment," in *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, pp. 435–447, 2017. DOI:10.1145/3116595.3116635.

[70] F. Baumann, D. Emmert, H. Baumgartl, and R. Buettner, "Hardcore gamer profiling: Results from an unsupervised learning approach to playing behavior on the steam platform," *Procedia Computer Science*, vol. 126, pp. 1289–1297, 2018. DOI:10.1016/j.procs.2018.08.078.

[71] B. Liu, *Sentiment Analysis and Opinion Mining*. San Rafael, California: Morgan & Claypool Publishers, 2012. DOI:10.1007/978-3-031-02145-9.

[72] F. Tang, L. Fu, B. Yao, and W. Xu, "Aspect based fine-grained sentiment analysis for online reviews," *Information Sciences*, vol. 488, pp. 190–204, 2019. DOI:10.1016/j.ins.2019.02.064.

[73] A. J. Viera, J. M. Garrett, *et al.*, "Understanding interobserver agreement: the kappa statistic," *Fam med*, vol. 37, no. 5, pp. 360–363, 2005. DOI:https://pubmed.ncbi.nlm.nih.gov/15883903/.

[74] E. Conroy, M. Kowal, A. J. Toth, and M. J. Campbell, "Boosting: Rank and skill deception in esports," *Entertainment Computing*, vol. 36, p. 100393, 2021. DOI:10.1016/j.entcom.2020.100393.

[75] J. H. Hyeong, K. J. Choi, J. Y. Lee, and T.-H. Pyo, "For whom does a game update? players' status-contingent gameplay on online games before and after an update," *Decision Support Systems*, vol. 139, p. 113423, 2020. DOI:10.1016/j.dss.2020.113423.

[76] H.-N. Kang, H.-R. Yong, and H.-S. Hwang, "A study of analyzing on online game reviews using a data mining approach: Steam community data," *International Journal of Innovation, Management and Technology*, vol. 8, no. 2, p. 90, 2017. DOI:10.18178/ijimt.2017.8.2.709.

[77] "2022 asian games," accessed Sep-2022. https://www.aesf.com/en/News-And-Media-Coverage/Esports-Returns-As-Proposed-By-The-Hagoc-For-The-2022-Asian-Games.html.

[78] "Global esports market in 2020," accessed Sep-2022. https://newzoo.com/insights/trend-reports/newzoo-global-esports-market-report-2020-light-version/.

[79] C. R. Aydin and T. Güngör, "Combination of recursive and recurrent neural networks for aspect-based sentiment analysis using inter-aspect relations," *IEEE Access*, vol. 8, pp. 77820–77832, 2020. DOI:10.1109/ACCESS.2020.2990306.

[80] N. Li, C.-Y. Chow, and J.-D. Zhang, "Seml: A semi-supervised multi-task learning framework for aspect-based sentiment analysis," *IEEE Access*, vol. 8, pp. 189287–189297, 2020. DOI:10.1109/ACCESS.2020.3031665.

[81] Z. Kastrati, A. S. Imran, and A. Kurti, "Weakly supervised framework for aspect-based sentiment analysis on students' reviews of moocs," *IEEE Access*, vol. 8, pp. 106799–106810, 2020. DOI:10.1109/ACCESS.2020.3000739.

[82] W. Li and B. Xu, "Aspect-based fashion recommendation with attention mechanism," *IEEE Access*, vol. 8, pp. 141814–141823, 2020. DOI:10.1109/ACCESS.2020.3013639.

[83] A. Agrawal, C. Silva, and P. Xenopoulos, "Data science and analytics for esports," in *2022 IEEE Integrated STEM Education Conference (ISEC)*, pp. 386–386, IEEE, 2022. DOI:10.1109/ISEC54952.2022.10025321.

[84] N. H. L. Wong, S.-i. Kwok, and J. L. Tan, "Predictive analytics of first blood and match outcome in dota 2," in *TENCON 2022-2022 IEEE Region 10 Conference (TENCON)*, pp. 1–6, IEEE, 2022. DOI:10.1109/TENCON55691.2022.9977786.

[85] A. Wattanaburanon and N. Propoon, "Method for classifying usability qualities and problems for action games from user reviews using text mining," in *2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS)*, pp. 1–6, IEEE, 2016. DOI:10.1109/ICIS.2016.7550748s.

[86] Y. Yu, D.-T. Dinh, B.-H. Nguyen, F. Yu, and V.-N. Huynh, "Mining insights from esports game reviews with an aspect-based sentiment analysis framework," *IEEE Access*, vol. 11, 2023. DOI:10.1109/ACCESS.2023.3285864.

[87] Y. Yu, B.-H. Nguyen, D.-T. Dinh, F. Yu, T. Fujinami, and V.-N. Huynh, "A topic modeling approach for exploring attraction of dark souls series reviews on steam," in *Intelligent Human Systems Integration (IHSI 2022): Integrating People and Intelligent Systems*, vol. 22, AHFE Open Acces, 2022.

[88] Y. Yu, T. Dinh, F. Yu, and V.-N. Huynh, "Understanding mobile game reviews through sentiment analysis: A case study of pubgm," in *International Conference on Model and Data Engineering*, pp. 102–115, Springer, 2023.

[89] T. Taylor, "Watch me play," in *Watch Me Play*, Princeton University Press, 2018.

[90] R. P. Griffiths, M. S. Eastin, and V. Cicchirillo, "Competitive video game play: An investigation of identification and competition," *Communication Research*, vol. 43, no. 4, pp. 468–486, 2016. DOI:10.1177/0093650214565895.

[91] J. Juul, *Half-real: Video games between real rules and fictional worlds.* MIT press, 2011. URL:https://mitpress.mit.edu/9780262516518/half-real/.

[92] J. L. Ba, J. R. Kiros, and G. E. Hinton, "Layer normalization," *arXiv preprint arXiv:1607.06450*, 2016.