UNIVERSITY OF PELOPONNESE
FACULTY OF HUMAN MOVEMENT AND QUALITY OF LIFE SCIENCES
DEPARTMENT OF SPORTS ORGANIZATION AND MANAGEMENT

THE IMPACT OF THE INCLUSION OF SKATEBOARDING IN THE OLYMPIC GAMES ON THE DEVELOPMENT OF SKATEBOARDING PROGRAMS——THE CASE OF CHINA

By Zehua Zhang

MASTER Dissertation submitted to the professorial body for the partial fulfillment of obligations for the awarding of a post-graduate title in the Post-graduate Programme, "Olympic Studies, Olympic Education, Organization and Management of Olympic Events" of the University of the Peloponnese.

Sparta 2024

Approved by the Professor body:

1st Supervisor: Kristine Toohey Emeritus Prof. GRIFFITH UNIVERSITY, AUSTRALIA

2nd Supervisor: Dimitris Gargalianos Prof. DEMOCRITUS UNIV. OF THRACE, GREECE

3rd Supervisor: Cesar R. Torres Prof. STATE UNIV. OF NEW YORK BROCKPORT, AMERICA

SUMMARY

Zhang Zehua: The impact of the inclusion of skateboarding in the Olympic Games on the development of skateboarding programs—— the case of China.

(Under the supervision of Kristine Toohey, Emeritus Professor)

The purpose of this study is to analyze the impact of the introduction of skateboarding to the Olympics Games on the development of skateboarding in The People's Republic of China. This paper analyzes the relevant comments on skateboarding on Chinese social media before and after skateboarding's inclusion in the Olympics Games using web python, the theory of life cycle, and sentiment analysis theory, in order to investigate resultant changes in public awareness and emotions. The Olympics attracts worldwide attention. The result of this paper provide suggestions for the long-term development of skateboarding. In the study, first collected data about skateboarding discussion from Weibo using web python technology and analyzed the data using sentiment analysis theory to study the change of the public's emotional tendency towards skateboarding. Then, analyzed the discussion trends of skateboarding in Weibo using the theory of life cycle to understand the change of public sentiment after skateboarding joined the Olympic Games from the perspective of the public. The results show that after skateboarding was included in the Olympics Games, the public's attention and positive emotions towards skateboarding increased. At the same time, through the analysis of the life cycle theory, I observed different stage characteristics in the discussion of skateboarding on Weibo, which is of great significance for understanding the development trajectory of skateboarding. Based on the comprehensive analysis results, this paper proposes some suggestions, including strengthening the promotion of skateboarding on social media, promoting the popularization of skateboarding culture, and improving the infrastructure of skateboarding projects. These suggestions may contribute to further promoting the development of skateboarding and provide references for the development of similar projects.

Keywords: Skateboarding, Olympic Games, emotional classification, network public opinion analysis, main body mining

ACKNOWLEDGEMENT

I would like to express my sincere gratitude towards my tutor Kristine Toohey, for her patient and careful guidance through whole year.

In addition, I would like to express my sincere gratitude to the other two outstanding professors on my thesis committee, Dimitris Gargalianos and Cesar R. Torres. Thank you very much for your trust and assistance.

I am deeply grateful to my parents. Thank you for their support, both financially and emotionally. Being able to study abroad is also due to the support of my parents. They have always supported my decisions, provided me with affirmation and encouragement, and have been my rock-solid harbor.

Thank you to every professor throughout the master's process. Your sense of humor, professional knowledge, and help have benefited me greatly.

Thank you to every classmate in class and young ambassador. It has been a fun and meaningful experience to make friends with people from different cultural backgrounds. I hope to meet you all again in any corner of the world in the future!

Thank you, Greece, and Mr. Pierre de Coubertin, the pioneer of the Olympic Games!

Thank myself!

Lastly, thank you to my great motherland China!

CONTENTS

LIST OF FIGURERS

LIST OF TABLES

CHAPTER I

INTRODUCTION

1.1 Introduction to Skateboarding Projects

Skateboarding, as an emerging sport, is beloved and admired by young people. It originated from street culture and is a unique experience that combines sports, art, and freedom. However, over time, skateboarding has complemented street culture, evolving into a highly respected competitive sport. Its growth has even led it to the Olympic stage, becoming an official event in the Olympic Games.

Skateboarding originated in California, United States in the 1950s (Lombard, 2015). It initially began as a way for surf enthusiasts to enjoy themselves on land when there were no waves. They attached roller skate equipment to a wooden board and imitated the feeling of surfing through actions such as sliding and jumping, thus giving birth to skateboarding. As time went on, this sport gradually became popular among young people in the United States and spread to various parts of the world.

Skateboarding is divided into multiple levels, including street skateboarding, park skateboarding, and vertical skateboarding, each with its unique techniques and competition rules. These competitions are not only held worldwide but also attract a large number of spectators and sponsors.

With the professionalization of skateboarding, more and more professional skateboarders have emerged. Through long periods of training and competitions, they have mastered various difficult techniques and have demonstrated amazing performances in competitions. Some well-known skateboarders such as Tony Hawk, and Nyjah Huston have become legendary figures in the skateboarding community. Their performances not only inspire more young people to join skateboarding but also bring more attention to the sport.

In 2016, the International Skateboarding Federation (ISF) successfully recommended skateboarding as an official event in the Olympic Games, and it made its first appearance at the 2020 Tokyo Olympics. The formal inclusion of skateboarding in the

Olympics signifies that the sport has transformed from a street culture into an internationally recognized competitive sport.

1.2 The Importance of New Olympic Sports

The Olympic Games, organized by the International Olympic Committee(IOC), is one of the largest and most influential comprehensive sports events in the world (Malfas et al., 2004). In 1894, Pierre de Coubertin, known as the ‘father of the modern Olympics,’ decided to establish the International Olympic Committee and initiate the modern Olympic movement. Since the first modern Olympic Games held in Athens, Greece in 1896, it has experienced nearly 130 years of ups and downs, hosting 32 editions. The first Olympic Games had a total of 9 major events, and the Tokyo Olympic Games in 2020 had 33 major events and 339 sub-events, reflecting the development trend of the Olympic movement (Jiao & Song, 2023). The Olympic Games demonstrate the various functions of sports, encourage more people to participate in sports to keep fit and feel happiness, Youth Olympics play the educational significance of sports, and their influence goes far beyond the scope of sports, producing a series of significant impacts in contemporary world politics, economy, philosophy, culture, art, and news media（Scheu et al., 2021). According to the 51st Statistical Report on the Development of China's Internet issued in March 2023, as of December 2022, the number of Chinese Internet users has increased by 35.49 million compared to December 2021, reaching 1.067 billion, and the popularity of the Internet has reached 75.6%. The proportion of Internet users accessing the Internet through mobile phones is as high as 99.6% (China Internet Network Information Center, 2023). With the widespread use of the Internet, the way information is disseminated has undergone significant changes. Compared to traditional paper-based information dissemination, the internet as a medium has broadened the channels for user comments and communication. People share their lives on social media platforms, express their opinions on social hot topics, and engage in mutual exchange of ideas, resulting in a surge in platform data. Internet users have also transformed from passive recipients of internet information to producers of information. These textual data often contain users' subjective emotions towards specific objects, and analyzing the emotional information in these data can bring significant benefits to governments, sports enterprises, and ordinary users(Zhou et al., 2019).

Olympic sports, as the pinnacle stage of athletics, have evolved and changed to adapt to the development of society and the interests of people (Garcia, 2008). As a road map for the reform of Olympic sports, the Olympic Agenda 2020 proposed that the Olympic Games should be more flexible, innovative, sustainable, and gender equal in terms of the selection of Olympic events (Thorpe & Wheaton, 2019). The inclusion of different sports in the Olympic Games will undoubtedly attract unprecedented attention and discussion, and will also lead to new trends and directions in sports. Among them, the introduction of new sports is considered one of the key factors for the continuous attraction of global audiences and athletes to Olympic sports(Hao, 2019). This study focuses on skateboarding as a new addition to the Olympic Games, exploring the changes in public attitude towards skateboarding after its inclusion.

With the continuous development of Olympic sports, the introduction of new sports not only expands the diversity of the Olympic Games but also provides a global platform for the newly included sports. This expansion not only has a profound impact on participating countries and athletes but also shapes the public's attitude and perception of different sports. The inclusion of skateboarding, as a highly fashionable and youthful sport, has attracted widespread attention (Beal et al., 2017).

Therefore, studying the performance of skateboarding on the Olympic stage and its impact on public perception has both theoretical and practical significance.

1.3 Theoretical Basis

With the inclusion of skateboarding as part of the Olympic Games, the sport's influence has gradually expanded globally and has garnered widespread attention. To better understand the change in public attitudes towards skateboarding and the impact of the Olympics on the development of the sport, a thorough analysis based on a series of theoretical frameworks is required. This study will comprehensively explore the attitude change and development impact of skateboarding after its inclusion in the Olympics, based on the theories of Olympic influence, sentiment analysis, and event analysis.

The literature reviewed for this study included materials investigating research of event network public opinion, the research on the evolution of public opinion, and research on sentiment and themes. This lead to the identification of existing problems in the research on the evolution of event network public opinion, and thus the use of a sentiment classification model and theme evolution analysis method for this study. The study then introduces the application of the lifecycle theory and cognitive affective appraisal theory, providing a theoretical basis and laying a solid foundation for the subsequent research. The study also explores how the addition of skateboarding as a new Olympic event shapes public perception.

This study will comprehensively utilize theories of the influence of the Olympics, emotional analysis methods, public opinion analysis methods, and event analysis methods to deeply analyze the changes in the public's attitude towards skateboarding after its inclusion in the Olympics and its impact on the development of skateboarding in the People Republic of China. Through the combination of theories, I hope to comprehensively interpret the profound impact of the Olympics as a global sporting event on skateboarding, providing beneficial insights for its future development on the international sports stage.

1.3.1 Theory of Olympic Influence

The Olympic Games, as one of the largest sporting events in the world, possess a unique and powerful influence. Its global platform for dissemination has made the Olympics a medium that leads sports development and changes public perceptions. The theory of Olympic influence emphasizes that this international event can promote sports to the world stage and influence people's cognition and attitudes through various channels such as media and social platforms (López et al., 2020). As a newly added Olympic event, skateboarding will be shaped by this theory of influence, thereby triggering a change in public attitudes towards the sport.

1.3.2 Emotional Analysis Method

The emotional analysis method plays an important role in studying the attitude changes of the public towards skateboarding projects (Wheaton et al., 2017). This method focuses on individuals' emotional experiences towards specific topics or events, revealing public preferences, attitudes, and emotional fluctuations through the analysis of language and emotional expressions on social media. By applying emotional analysis, it is possible to can gain a deep understanding of the emotional responses triggered by skateboarding joining and officially appearing in the Olympic Games, thereby interpreting the public's attitude changes towards this project.

1.3.3 Event Analysis Method

The event analysis method emphasizes the profound impact of large-scale events that occur at specific times and places on society (Coalter & Taylor, 2009). As a global mega-event, the Olympics not only has a significant impact on the host city but also attracts worldwide attention. Applying the event analysis method to the context of skateboarding joining the Olympics helps to understand the driving force of the Olympics for the development of skateboarding and its enduring impact on the public.

1.4 Research Questions

This study will focus on the following questions: RQ1: What changes have occurred in public perception of skateboarding as an Olympic sport? RQ2: The peak period of emotional changes is at what stage? RQ3: What kind of public opinion atmosphere has arisen as a result of these changes? RQ4: How has the Olympic platform influenced the recognition and acceptance of skateboarding among the Chinese population?

1.5 Hypothesis

Based on the theories of the lifecycle theory, the emotional analysis theory, and the universal trend of adding new events to the Olympics, I assume that the introduction of skateboarding as an Olympic event will lead to positive changes in public perception of this sport. In this study, I will comprehensively explore these issues through event analysis, public opinion analysis, and sentiment analysis, among other methods, in order to provide new perspectives and insights for future studies of new Olympic events. By gaining a deeper understanding of public perception of skateboarding, I hope to provide valuable insights for the future development of this sport and the evolution of Olympic sports.

1.6 Limitation

This article focuses on the selection of comments regarding the inclusion of skateboarding in the Olympic Games. The comments analyzed are mainly from active youth users of Chinese social media platforms, and the sample does not represent the opinions of all internet users. Future research should consider using data from different types of media platforms or other techniques to obtain a more comprehensive understanding of public opinion. Additionally, this study only investigates skateboarding as a new addition to the Olympic Games, and it is unknown whether there are differences in public opinion between skateboarding and other sports.

CHAPTER Ⅱ

LITERATURE REVIVEW

This chapter presents relevant research conducted on the thesis topic. It also provides a brief introduction to the relevant theories used in the study's analysis specifically regarding the emotional evolution and thematic characteristics analysis of the network public opinion events.

2.1 Conceptual Research on Network Public Opinion Events

Public opinion, as a concept that has developed in the Chinese context, was initially understood by Chinese scholars as the social and political attitudes of the public. In a certain social environment, the concept of public opinion has a profound influence on the values, political orientations, and social attitudes of the public as social actors, revolving around mediated social events. In the specific Chinese context, public opinion refers to the values, emotional expressions, and behavioral tendencies of the public (including individuals and various social groups) towards a certain type of event(Tao, 2014). In the era of new media, the expansive dissemination of the Internet, mobile phones, and other self-media has broadened the channels and methods through which the public obtains information, optimized the pathways of information dissemination, intensified the dissemination of public opinion information, and expanded the scope of dissemination, increased the speed of dissemination, and complicated the pathways of dissemination, resulting in an increase in the intensity and breadth of public opinion dissemination compared to the old media era (Tang, 2015). Network public opinion, on the other hand, emerged based on the development of the Internet and represents a particular type within social public opinion(Liu, 2012). Currently, scholars have not yet reached a consensus on the definition of network public opinion. Ji (2007) studied network public opinion and defined it as the social and political views held by netizens. Li et al. (2015), on the other hand, regarded network public opinion as the sum of the various emotions, attitudes, and understanding displayed by people after a certain event. In addition, Zeng(2014), Zhou(2016), and others defined network public opinion as the expression of the public's opinions and insights on public events with the help of the media on online platforms. Feng et al. (2012) conducted research and pointed out that in the era of new media, the dissemination of public opinion is characterized by rapid speed, wide breadth, difficulty in control, intensified interaction and influence of information exchange, difficulty in verifying the source, and a greater prevalence of negative emotions. Li et al. (2014) also conducted research and identified characteristics of public opinion dissemination such as the generalization and popularization of subjects, the complexity of content, the interaction in the dissemination process, the rapid formation of public opinion, the diversification of the dissemination process, the uncontrollability of dissemination, and the polarization of public opinion groups. In summary, as a form of public expression of social situations, public opinion dissemination rapidly spreads among various social groups through modern communication means, exhibiting its own characteristics during the dissemination process.

Sports events, as significant social phenomena, have a profound impact on the local economy and society (He & Li, 2008). The organization of sports events is influenced to a certain extent by the social environment in which they take place. Sports events receive high levels of attention. For example, the live broadcast of the opening ceremony of the Beijing Olympics was watched by up to 1 billion people, creating the highest viewership in human history. The sixth game of the 2013 NBA Finals had a viewership rating of 14.7 in North America, making it the highest-rated program in the United States that year (Foster et al., 2014). The high level of attention and the dynamic interaction between sports events and politics, society, and the economy intensify the attention received by sports events. The speed and breadth of public opinion dissemination about sports events also increase in line with the level and breadth of attention received (Xie et al., 2022). Based on research on online public opinion, Jin (2014) believes that the analysis of online public opinion is increasingly mature and has been successfully applied in related fields, such as online crisis public relations, social hot spot tracking, marketing activities effect monitoring and advertising media research and so on. But its application in the field of sports is still in its infancy. Sun et al. (2017) argue that the dissemination of sports information on the internet exhibits strong characteristics, showing a strong tendency towards youthfulness and using intuitive and concise forms of expression. In the era of new media, public opinion, as a form of social network communication, spreads along social networks, and the dissemination of sports events follows the same pattern. The dissemination of public opinion about sports events is characterized by responsibility, interactivity, and variability, and is accompanied by fast dissemination and high levels of attention. While sports events receive attention, the changes in sports events, especially in events like the Olympic Games, are also closely watched (Fan, 2019).

Based on the understanding and definition of the concept of network public opinion about events by scholars mentioned above, this article accepts that network public opinion about events refers to the sum of the media and the public expressing their views, opinions, and emotions about the events through social media platforms such as Weibo and WeChat. In the context of this thesis, social media and the inclusion of skateboarding as an official Olympic event and its first appearance in the Olympics serve as the context of the study.

2.2 Lifecycle Theory

The theory of the life cycle is involved in many fields such as business, economy, society, and environment. It is a theory that integrates multiple disciplines and domains. It originated in the field of biology, which believes that the life cycle is an objective law followed by all organisms in the natural world. It reveals the repetitive process that every individual in nature will go through, from formation to eventual extinction. The original life cycle theory in the field of biology refers to the individual meaning, but it has now been expanded and extended to other new fields, such as product life cycle, demand life cycle, and so on (Zhao, 2003).

Sudden online public opinion events also have a certain life cycle, that is, any specific event will follow the law of formation and eventual decline (Liu, 2019). Therefore, many scholars, in order to study the characteristics of stage changes in the network public opinion of sudden events, have divided the entire process of sudden events into three stages or four stages, or even more stages, using the life cycle theory. Du et al. (2017) used a three-stage model to explore the evolutionary patterns of network public opinion in different sudden events. Hou and Zhang (2018) studied the dissemination mechanism of negative public opinion information in sudden events, mainly including the stages of opinion brewing, dissemination, change, and decline. In addition, Cui et al. (2018) divided the development cycle of network public opinion into six periods: emergence, outbreak, dissemination, diffusion, recurrence, decline, and long tail. Wang (2023) analyzed the mechanism of public opinion guidance for sudden events in the era of digital intelligence from the perspective of sports events, and it can be divided into five operational levels: monitoring and early warning mechanism, judgment mechanism, response mechanism, evaluation mechanism, and communication feedback mechanism. Kwon & Oh (2019) have conducted exploratory research on the life cycle of international sports events, and Shin (2007) identified boxing as a "product" and used the product life cycle theory to analyze changes in the popularity of boxing.

2.3 Cognitive-Affective Appraisal Theory

Emotion is an individual's attitude towards something that satisfies their needs (Du, 2014). With the development of computer applications, people have become aware of the limitations of computer technology in understanding emotions. To address this issue, scholars have integrated cognitive psychology and computer science, using rule-based language to provide a scientific theoretical basis for emotion evaluation in the field of computer science. Currently, researchers widely use the OCC model to create rule-based models of emotional cognition (Adam, Herzig, & Longin, 2009).

Ortony, Clore, and Collins (1988) proposed the Cognitive Appraisal Model of Emotions, known as the OCC model. The OCC model, named after Ortony, Clore, and Collins presented their 1988 book "The Cognitive Structure of Emotions," is a cognitive-emotional evaluation model. It classifies human emotions into twenty-two types based on three major dimensions: Consequences of Events, Action of Agents, and Aspects of Objects. According to the OCC model, humans have 22 different emotions, each generated through different paths based on criteria, triggers, and intensity dimensions. The model presents a comprehensive and detailed tree structure, branching out various emotions based on the outcomes of events, the image of objects, and the behavior of objects. It has been widely cited by researchers both domestically and internationally. Currently, the OCC model is mainly applied in theoretical logic formalization, artificial intelligence, and e-education fields, with limited use in text processing(Huangpu & Mao, 2017). Shaikh et al. (2006) demonstrated the feasibility of using the cognitive appraisal model for sentiment classification of text. Pang (2008) developed a model for text sentiment recognition based on the OCC cognitive appraisal model, and comparative experiments showed that the OCC text sentiment recognition model outperforms other models due to the consideration of cognitive appraisal elements.

Here are some examples of emotions and their triggers based on the OCC model: Happiness: Generated when an event or object is appraised as relevant to one's goals and conducive to achieving them. For example, receiving praise for a job well done can trigger feelings of happiness. Sadness: Arises when an event or object is appraised as relevant to one's goals but not conducive to achieving them. For instance, the loss of a loved one can evoke feelings of sadness. Anger: Triggered by appraisals of an event or object as obstructing or thwarting one's goals. For example, being unfairly criticized or treated can lead to feelings of anger. Fear: Generated when an event or object is appraised as posing a threat or danger. For instance, encountering a wild animal in the wilderness can evoke feelings of fear. Surprise: Arises when an event or object is appraised as unexpected or incongruent with one's expectations. For example, receiving an unexpected gift can trigger feelings of surprise. Disgust: Triggered by appraisals of an event or object as offensive or repulsive. For instance, encountering spoiled food can evoke feelings of disgust.These examples illustrate how different emotions can be generated through cognitive appraisals of various events, objects, and behaviors, as proposed by the OCC model(Luo, 2023).

However, there is limited research on using the OCC model for online public opinion detection in China. While there are some studies in this area, they are relatively few in number and scope compared to the overall research landscape. This study aims to combine the OCC model with deep learning model to describe the emotional states of netizens' Weibo posts, providing a new approach and method for public opinion detection(Liu, 2018).

CHAPTER III

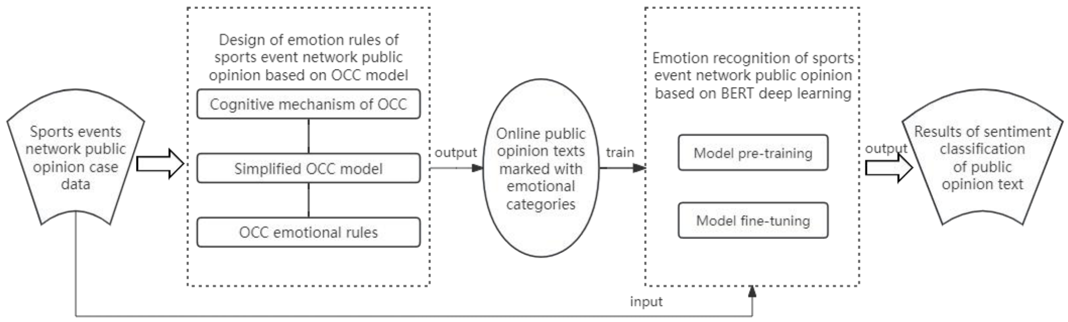
METHODOLOGY

This chapter explains the methodology used in this study. It explores the sentiment analysis of public opinion texts in event networks, which refers to the collective expressions of emotions and opinions by netizens on social media platforms regarding relevant events. Therefore, sentiment identification of public opinion texts is one of the key focuses in studying event network public opinion. This chapter will construct an event network public opinion sentiment classification model based on Ortony, Clore, and Collins model & Bidirectional Encoder Representations from Transformers (OCC-BERT). The sentiment classification model mainly consists of OCC cognitive emotion rules and BERT deep learning algorithm. This chapter will elaborate on the implementation process of these two methods and verify the performance of this sentiment classification model through experiments.

3.1 Model construction

Based on the current domestic and international research results on emotion classification, commonly used emotion classification methods include sentiment lexicon method, machine learning method, and deep learning method. Machine learning method and deep learning method provide many classification algorithms with high accuracy, but to some extent, they rely on the accuracy of the emotion category labels in the dataset. In the past, emotion category labels in the dataset were generally annotated manually, and this annotation method has a large degree of subjectivity and lacks theoretical support, which may lead to a decrease in the accuracy of emotion classification models (You et al., 2022). In order to solve this problem, this chapter introduces methods from the field of psychology and formulates OCC emotion classification rules based on the OCC model cognitive framework, which are used to annotate the emotion categories of textual data on sudden events in online public opinion, in order to reduce the subjectivity of emotion judgment to some extent. At the same time, considering the huge scale of textual data on sudden events in online public opinion and the low efficiency of manual annotation, the Bidirectional Encoder Representations from Transformers(BERT) model in deep learning technology is referenced to automatically identify emotional categories of textual data on sudden events in online public opinion. BERT is a pre-training language model developed by Google, which can be used to generate and process natural language texts. In the BERT model, both the input and output are a set of pre-trained word sequences, and the core encoding architecture is a multi-layer bidirectional Transformer(Li et al, 2023). This paper applies the pre-training model and fine-tuning steps of BERT to the sentiment classification of event network public opinion. This study collected and obtained textual data on sudden events in online public opinion, and labeled the emotional data according to OCC rules, outputted three categories of positive, negative, and neutral emotional labels of textual data on online public opinion as the training set for pre-training and fine-tuning in the BERT model. Then, the collected textual data on sudden events in online public opinion was input into the trained OCC-BERT model to recognize the emotions of the predicted textual data, and finally outputted the emotion classification results of a large-scale textual data. The specific methodology stages are shown in Figure 1.

Figure 1. Research roadmap



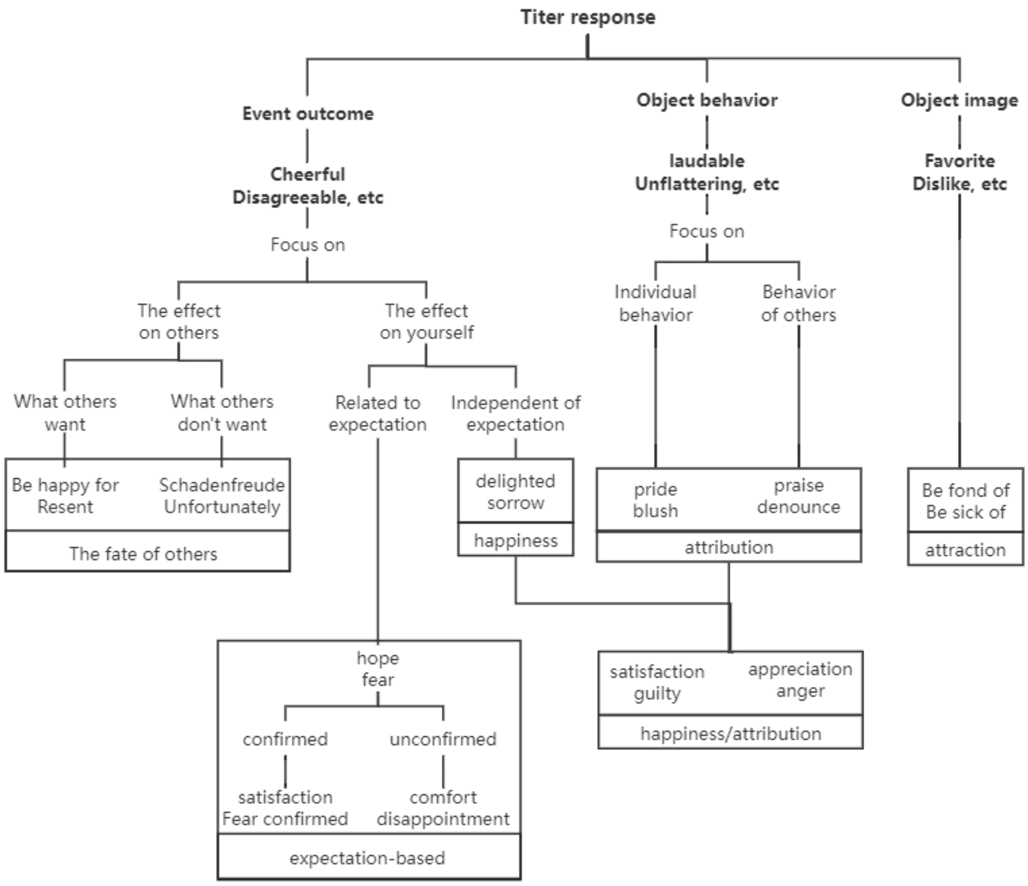
3.2 OCC cognitive emotion evaluation model

3.2.1 Overview of the OCC model.

Emotions play a role in the process of interaction between individuals, serving as a means of expressing opinions. People need to explore emotions and express opinions through the use of emotional models (De et al., 2012). The OCC model, named after

Ortony, Clore, and Collins presented their 1988 book "The Cognitive Structure of Emotions," is a cognitive-emotional evaluation model. It classifies human emotions into twenty-two types based on three major dimensions: Consequences of Events, Action of Agents, and Aspects of Objects. The first category focuses on whether the outcome of an event is pleasant or not. The second category emphasizes whether the behavior of the object in the event is satisfactory. The third category focuses on whether the object itself in the event is likable. Different emotions within each category are further differentiated and categorized into twenty-two different emotional states based on their intensity and causes. The structure of the OCC model is shown in Figure 2(Ortony, Clore, & Collins, 1988).

Figure 2. Classic OCC model structure diagram



In the OCC model, emotions are the result of a cognitive judgment process. The generation process is as follows:

(1)Classification: Evaluation in the OCC model relies on three components: consequences of events; action of agents; and aspects of objects. It is necessary to classify them initially in order to evaluate the emotional impact of three evaluation criteria on the information recipient. In the process of evaluation, if the focus is on the consequences of events, the target is the main concern. If the evaluator attaches more importance to the behavior of the object in the event, the behavior criterion is the main concern. If the evaluator focuses on the image of the object, then the attitude towards the evaluation object itself, whether it is love, hate, sadness, or resentment, is the main concern (Ortony, Clore, & Collins, 1988).

(2) Quantification: Consideration is given to whether the information recipient's acceptance of the information is strong enough and whether it will have an impact on their emotional state (Yang, 2018).

(3) Mapping: As shown in Figure 2., the OCC model classifies human emotions into twenty-two types through the three main evaluation dimensions in the classification process and a series of subsequent emotional triggering conditions. These emotions include "Love," "Joy," "Distress," "Shame," and so on. Each emotion type has its own unique generation path. Taking the emotion of "Shame" as an example, its generation path is as follows: the emotional dimension is a reaction to the behavior of the object - the behavior of the object is not satisfactory and praiseworthy - attention is focused on oneself - negative evaluation of behavior - "Shame." All twenty-two emotions can be mapped to fewer dimensions, simplifying the mapping and categorizing all emotions into six types. Emotions can also be mapped into positive and negative polarities(Dong, 2008). This study simplifies the mapping of the emotional model and categorizes the sentiment of public opinion comments on sudden events into positive sentiment, negative sentiment, and neutral sentiment.

(4) Expression: The emotions generated by the aforementioned series of triggers exist in the minds of the emotion generators and need to be manifested through expression methods. Typically, facial expressions, body language, and words are used for expression. Different emotions have their own unique ways and means of expression. For example, someone may express a happy mood with a relaxed and pleasant tone

while smiling and speaking. Comments posted online can also express one's emotions and viewpoints. For example, when discussing a hot topic on a forum, netizens will voice their opinions to express their different views (Yao, 2009). In this study, the sentiment of public opinion on sudden sport events is analyzed through online social platforms. Therefore, facial expressions and body language of evaluators cannot be obtained, but emotions are identified through the content and tone of the texts posted by netizens. In order to make the sentiment labeling system of online public opinion "reasonable and standardized," this research refers to the aforementioned OCC emotional model with a rational and clear structure.

In order to better adapt to the emotional classification of netizens during sport events, the OCC model is simplified and emotional rules are established based on the actual background of public opinion on sudden events, as explained below.

3.2.2 Simplified OCC model

According to the OCC model people's emotional states can be classified into 22 different types, however they may feel confused when trying to distinguish all emotions in detail, making it difficult to apply the model to sudden event network public opinion. This study focuses on the text of sudden event network public opinion on Weibo platform, so for this thesis the OCC model was simplified based on the actual situation of event network public opinion Weibo text, in order to apply it to the emotional classification of netizens in sudden event network public opinion. The following paragraph explains the analysis of sudden event network public opinion text combined with specific situations:

(1) When the public expresses their views and attitudes towards a particular event online, significant differences among individuals can be observed. Due to diverse upbringing environments, they hold different sets of values, leading to varying evaluations of the same event based on their respective standards(Yang, 2014).

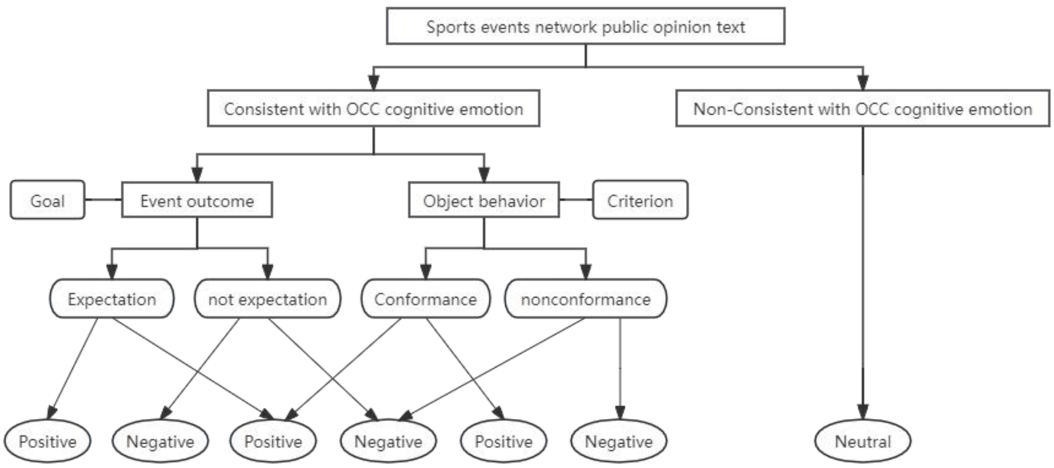
(2) The public has anonymity in the online environment, and when netizens deal with sudden events in a situation where they do not know each other, they tend to take a third-party perspective, so their opinions are more directed towards others rather than themselves (Zheng, 2011).

(3) Topics in the corpus related to events contain more negative emotions than other related events(Huang et al., 2015).

(4) No evaluation of the object itself. In event network public opinion, netizens focus on the event itself and the behavior of the objects involved, and they do not pay attention to or evaluate the image of the object itself(Zhang & Luo, 2021).

Based on the above analysis of the features of event network public opinion text, this study simplifies the OCC model that includes complex emotions and maps it into three categories: positive, negative, and neutral emotions, as shown in Figure 3., to adapt to the construction of OCC model for sports event network public opinion.

Figure 3. OCC affective cognitive model of online public opinion of sports events



This thesis eliminates the emotional evaluation dimension of object image in the three major dimensions, because it was found in the above analysis that the attitudes and opinions expressed by the public on the Internet are almost all from the perspective of bystanders, only evaluating the outcomes of the events or the actions of the objects. Therefore, the analysis is only conducted from the dimensions of event outcomes and object behavior. The first dimension is the focus on event outcomes, where the goal of the commentators is important, that is, considering how the commentators perceive the outcomes mentioned in the comment text. If they perceive that the event outcomes meet their expectations, positive emotions are generated. If the event outcomes do not meet their expectations, negative emotions are generated. The second dimension is to judge emotions based on the evaluation of the behavior of the objects. The principle of judgment is based on whether the behavior meets the standards or not. If the object's behavior complies with the behavioral norms, positive emotions are generated; otherwise, negative emotions are generated. Finally, emotions that cannot be inferred by the OCC model are classified as neutral emotions.

3.2.3 Emotion Rules of the Simplified OCC Model

According to the OCC emotion cognition model of sudden event online public opinion discussed in the previous section, this part elaborates on the simplified emotion rules of the OCC model to determine the polarity of emotions.

From the perspective of the OCC cognitive emotion model, there are two major judgment dimensions, namely the outcomes of sudden events and the behaviors that occur in sudden events. These dimensions include variables such as public opinion texts, events, behaviors, and emotions. Here, the relevant public opinion texts related to the research on sudden events are represented by "t", with a total of "i" text data. Each text data contains events that occur around this sudden topic, denoted as "e". In addition to discussing the outcomes of the events, each text data also discusses the behaviors of the objects in the text, represented by "a" (Ma, 2023).

To evaluate emotions from the perspective of event outcomes, the focus is mainly on whether the events meet expectations. The function formula is as follows:

$$Des(t_i, e) = Consequences(t_i, e) - Goals(e)$$

$Des(t_i,e)$ represents the degree to which the result of event e in the i-th text data t satisfies the public on the network platform. It is obtained by comparing the consequences of event e ($Consequences(t_i,e)$) with the satisfaction level towards the consequences of event e ($Goals(e)$). It can take two values, "less than 0" and "greater than 0". When $Des(t_i,e)$ is less than 0, it indicates that the result of event disposal in the text data does not meet the public satisfaction threshold, meaning that the event result does not meet the public's expectations. When $Des(t_i,e)$ is greater than 0, it indicates that the event result has reached the public's expected level.

Similarly, evaluating emotions from the perspective of object behavior in events mainly considers whether the behavior complies with the standards. The function formula is as follows:

$$Wor\,(t_i, a) = Actions(t_i, a) - S\tan dards\,(a)$$

Wor(t_i,a) represents the degree to which the result of event a in the i-th text data t satisfies the public on the network platform. It is obtained by comparing the object behavior in the i-th text (*Action(t_i,a)*) with the degree of agreement towards the object behavior (*S*tan*dard(a)*). It also has two values, "less than 0" and "greater than 0". When *Wor(t_i,a)* is less than 0, it indicates that the event disposal result in the text data does not meet the public satisfaction threshold, meaning that the public believes that the object behavior does not comply with the criteria. When *Wor(t_i,a)* is greater than 0, it indicates that the object behavior complies with the criteria perceived by the public. Combining *Des(t_i,e)* and *Wor(t_i,a)*, the function rule formula for judging the emotional category of each text data is as follows:

$$Emotions(t_i, e, a) = f\,(Des(t_i, e) - Wor\,(t_i, a))$$

Emotions(t_i,e,a) represents the emotional discrimination of each public opinion text. Its value range is [-1, 1, 0]. If its value is -1, it means that the emotional category of the public opinion text is negative. If it is 1, it represents positive emotion, and 0 represents neutral emotion. In the formula, f( ) represents that if *Des(t_i,e)* and *Wor(t_i,a)* occur separately and both are greater than 0, or if they occur simultaneously and both are greater than 0, it returns 1, representing positive emotion. If *Des(t_i,e)* and *Wor(t_i,a)* occur separately and both are less than 0, or if they occur simultaneously and both are less than 0, it returns -1, representing negative emotion. Otherwise, it returns 0, representing neutral emotion.

3.3 Emotion classification algorithm based on OCC-BERT

Based on the sentiment rules of the previous section of OCC, the sentiment annotation of sports event network public opinion text has a theoretical basis, which can obtain a high-quality annotation dataset and input it into the sentiment classification model. In the existing sentiment classification models, deep learning models have been widely welcomed by many researchers and have provided a good solution for researchers to deal with large-scale data (Luo, 2020). This study used the BERT deep learning algorithm to automatically identify the emotional categories of sudden event public opinion texts(Wan et al., 2024).

The pre-training task of BERT is essentially completed after unsupervised learning a large corpus, and its fine-tuning step uses the supervised learning of the sentiment annotation dataset of sports event network public opinion to adjust the parameters to meet downstream practical application requirements. Compared with other language models, on the one hand, the BERT model can better learn the positional relationship of word sequences(Zheng & Liu, 2023). Due to the non-standard comments in sports events and the possibility of ironic meanings in words, a word may have completely different or completely opposite meanings under changes in position and context. BERT improves this problem to a certain extent by adding position encoding information to the model. On the other hand, BERT can better learn the contextual relationship features of words because it adopts a multi-layer bidirectional Transformer structure, which enables it to better understand the semantics before and after the words instead of considering only one direction (Devlin et al., 2018).

This section will introduce the sentiment classification model from three parts: the basic structure of the BERT model, BERT pre-training and fine-tuning, and model design.

3.3.1 Basic structure of BERT model

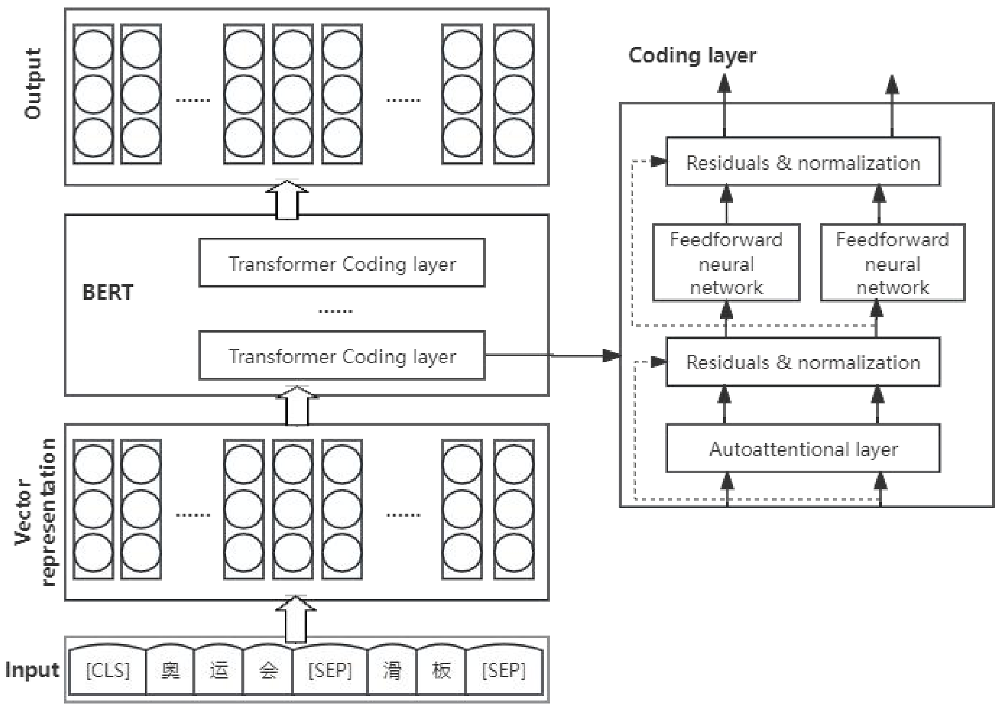
This study adopts the BERT-base Chinese model with a deep bidirectional Transformer architecture. Firstly, the sports event public opinion text data is inputted into BERT for embedding vector representation processing. Then, the Transformer core encoding layer is processed, which includes two sub-layers in multiple encoding layers, all of which are learned through feature extraction of the input signal and connected through hierarchical residual connections and normalization. One is the self-attention layer, which is a network based on attention mechanism that extracts important features from the input signal, thereby improving the training efficiency of the network. The other is the feed-forward neural network, which is a network based on connection structure that transmits useful information from the input signal to each output node in a nonlinear way, thereby improving the learning efficiency of the network. The output of the Transformer encoding is the final word embedding result (Vaswani et al., 2017). The structure of the BERT model is shown in Figure 4.

(1) Input Layer

When a public opinion text is inputted into the BERT model, the [CLS] mark is added at the beginning of the sentence, and the [SEP] mark is added at the end of the sentence. The text can be a sentence or a paragraph. In the case of a paragraph, the [SEP] mark is used to divide the two sentences, and each sentence is split into individual characters and inputted to the next layer.

Figure 4. BERT model structure diagram

奥运会 means Olympic Game; 滑板 means skateboarding



(2) Embedding Vector Representation Layer

The embedding text vector representation layer is the result of summing the three arrays of sentence layer, word vector layer, and position encoding layer. The first stage is going through the word vector layer, where the special tokens for the beginning and end of the sentence are inserted into the words, and then sent into the word vector layer, where each word is transformed into a vector of fixed dimensions, which can be used for subsequent classification tasks. Then, there is the sentence layer and the position encoding layer. The sentence layer distinguishes each sentence in a paragraph containing multiple sentences by recognizing the sequence mark, which can determine whether two sentence texts are similar. The position encoding layer combines the input word information with the position information, encoding the position of the words and enabling the model to learn word order.

(3) Transformer Layer

The Transformer layer is the core encoding layer of BERT, which consists of self-attention and feed-forward neural network. The calculation process of self-attention value is as follows:

① Establish the matrix $X_{N \times k}$ of the input text. Split the input text S and represent it as $S=\{w_1,w_2,\dots,w_i,\dots,w_n\}$, where w_i represents a word, i represents the position of the word in the input text, and n represents the length of the input text. After splitting, the input text can be represented as an input matrix $X_{N \times k}$. Each row of the matrix corresponds to the vector of a word in the input text, where k represents the dimension of the word vector.

② Construct the attention input matrices Q, K, V. Each matrix is obtained by multiplying the randomly initialized matrix W with the input matrix X, represented as: $Q=W^Q X^T$, $K=W^K X^T$, $V=W^V X^T$.

③ Calculate the attention value. The attention mechanism used by BERT is based on the scaled dot-product attention mechanism, which scales the attention value with a scaling factor $\sqrt{d_k}$. The attention value formula is as follows:

$$Attention(Q,K,V) = soft\,max\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$

BERT concatenates the attention values of 8 heads and multiplies them with the weight matrix W to obtain the multi-head attention values:

$$MultiHead(Q,K,V) = Concat(head_1,\dots,head_h)W^0,$$

$$head_i = Attention(QW_i^Q, KW_i^K, VW_i^V)。$$

where Concat( ) represents the concatenation function, W^0 is the weight matrix, W_i^Q, W_i^K, W_i^V respectively represent the weight matrices in the i-th head, and their dimensions are set the same as Q, K, V mentioned above.

(4) Vector Output Layer

After going through the data input layer, embedding layer, and Transformer layer, the final output is the vector representation of each word in the public opinion text, which integrates the semantic information of the entire text. It belongs to dynamic word vector representation.

3.3.2 BERT pre-training and fine-tuning

(1) BERT pre-training. BERT's pre-training consists of two parallel tasks: Masked Language Model (MLM) and Next Sentence Prediction (NSP). The process of MLM and NSP tasks will be explained below.

In the MLM task, the model is given a piece of text data during training, where one or more words in the text are randomly masked. The model is then required to predict the masked content based on the remaining unmasked text information. Training the MLM task introduces uncertainty when the BERT model encounters gapped positions, forcing the model to rely more on contextual information for word prediction and providing the model with error-correction capabilities.

The NSP task aims to train a model that understands sentence relationships. During model training, two sentences from a text data are provided, and the model is required to determine whether these two sentences are adjacent in the given text. In the actual pre-training process, 50 percent of correct and incorrect sentence pairs are randomly selected from the text dataset for training. Combined with the MLM task, this allows the model to better represent semantic information in the text.

The BERT model is trained jointly with the MLM and NSP tasks, ensuring that the output vectors of each model accurately reflect the overall information in the input text, providing better initial model parameter values for subsequent fine-tuning. In this study, I utilized a Chinese pre-trained BERT-base model and trained it with the default configuration. I then fine-tuned BERT on the annotated data obtained from scraping the burst event network public opinion text.

(2) Fine-tuning BERT. After completing the pre-training of BERT, fine-tuning for specific downstream applications can be done according to the task at hand. BERT can be applied to eleven natural language processing tasks in four major NLP domains, and has achieved excellent results. Different fine-tuning methods are employed for different tasks using BERT, including next sentence prediction, text classification, question answering, and word labeling tasks(Ye et al., 2023).

Since this study focuses on sentiment classification of comments in event network public opinion, which falls under the category of text classification tasks, the fine-tuning method for text classification is chosen. This method involves inputting a single sentence, adding the [CLS] separator at the beginning of the sentence, and determining the class label based on the output. This fine-tuning method is commonly used for single sentence tasks (Devlin et al., 2018).

In addition, fine-tuning can also be done for batch size, learning rate, and training epochs. Batch size can be adjusted based on hardware conditions, with larger GPU memory allowing for larger batch sizes. Adjusting the learning rate assigns a lower learning rate to the lower layers of BERT, which effectively fine-tunes the model. The number of epochs can be set to stop training if the training loss does not further decrease after a certain number of iterations(Yuan, 2021). The specific settings for these parameters will be detailed in the section on model parameter settings.

3.3.3 Model design

To enable the model to predict the sentiment category of sudden event network public opinion text, a Softmax linear classifier was added behind the BERT model that was fine-tuned with the emotion-labeled dataset. Therefore, the designed model for sentiment classification of sudden event network public opinion consists of three parts: public opinion data input basic training of BERT and linear output of sentiment categories. The sentiment classification model first takes in the sudden event network public opinion text data, then goes through multiple transformer encoding layers of BERT to output a feature vector H that contains contextual information. Finally, in the output layer, the learned semantic feature vector H is mapped to three categories:

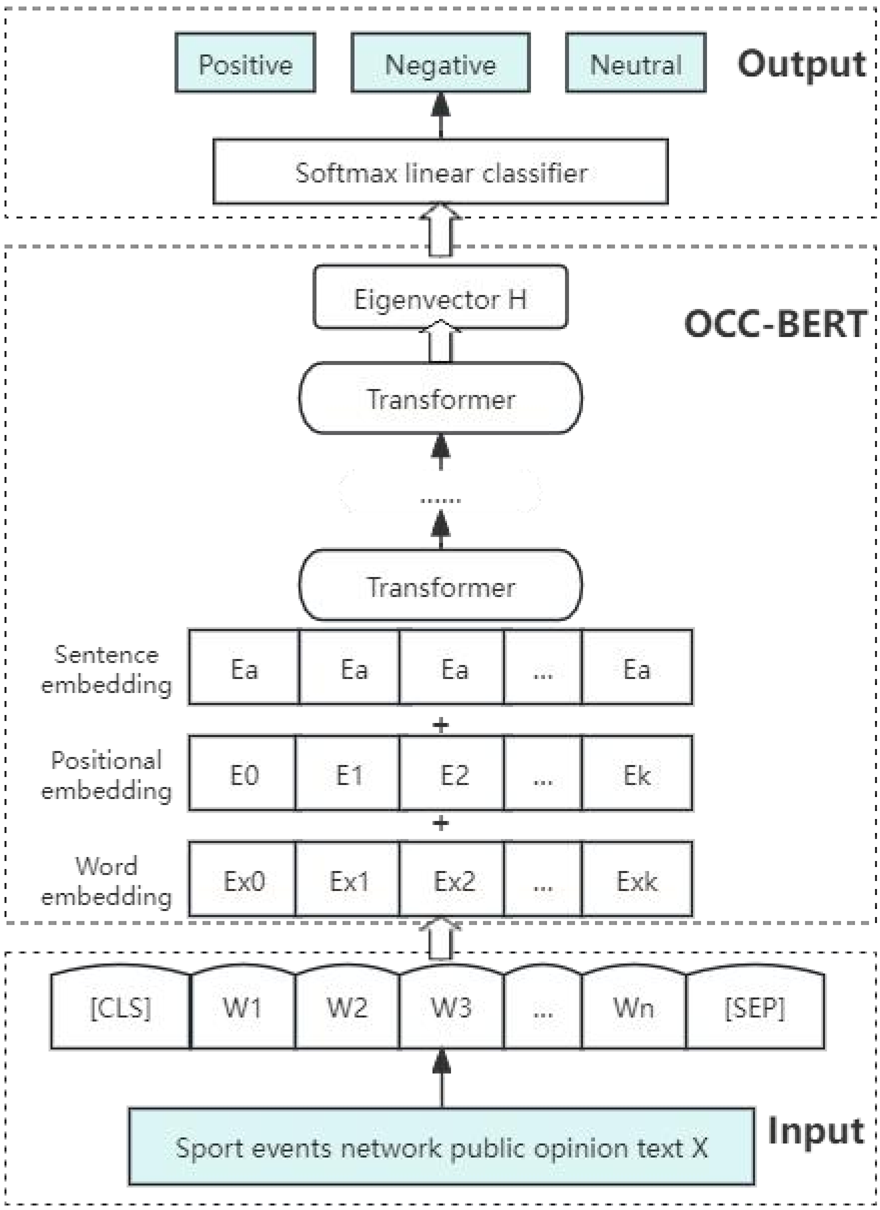
positive, negative, and neutral, using the Softmax linear classifier. The calculation formula is as follows:

$$P = Soft\max(W_h H + b_h)$$

Where W_h represents the weight matrix and b_h represents the bias.

The process of constructing the sentiment classification model for sudden event network public opinion is shown in Figure 5.

Figure 5. Sports events network public opinion emotion classification model flow



From the figure, it can be observed that the specific steps of the OCC-BERT sentiment classification algorithm are as follows:

(1) Prepare well-collected and processed sudden event network public opinion text data. Here, the input comment text is represented as a sequence $X = \{x_0, x_1, ..., x_k\}$;

(2) Input the raw text data into the trained OCC-BERT model, which decomposes the text data from sentence or paragraph form into individual characters. To distinguish between different sentences, the model adds the CLS and SEP identifiers at the beginning and end of each sentence, and then inputs them into the Embedding layer;

(3) After the data is input into the Embedding layer, it needs to go through the encoding of three sublayers: token embeddings, position embeddings, and segment embeddings;

(4) Then, through the multi-layer bidirectional transformer encoding structure, the semantic features and contextual information in the text data are learned, resulting in a feature vector H;

(5) After obtaining the extracted feature vector, it is then connected to the added fully connected layer. The function softmax is used to classify the sentiment of the sudden event network public opinion text, and obtain three sentiment categories: positive, negative, and neutral.

3.4 Experimental design and verification

This section explains the validation of the emotion classification model proposed in this chapter. First, in terms of dataset acquisition, the public opinion dataset about skateboarding joining the Olympics was collected using Python on the Weibo platform. Then, the experimental data was annotated with OCC method and non-OCC method to determine the sentiment categories. Next, the model parameters were set, and the experimental evaluation metrics are introduced. By comparing with the benchmark model, it was proven that using OCC sentiment rule annotated dataset improved the classification performance, and BERT method has superiority over other commonly used sentiment classification methods.

3.4.1 Experimental dataset

To train the emotion classification model, a certain amount of labeled text is needed as the training set of the model, and the trained model needs to be tested to directly verify its performance in the text domain(Wei, 2024). At the same time, in order to verify the effectiveness of OCC annotation, two types of labeled datasets are set in this experiment. Dataset: The public opinion comment data on the network of the current experiment is crawled from Weibo through a web crawler program. Taking "skateboarding joining the Olympics" as an example, the user comments on this sudden event are crawled. After filtering and cleaning the crawled data, there are a total of 11,251 valid data. Then, two types of experimental datasets are annotated based on natural annotation and OCC sentiment rule theory respectively. One is the natural annotation case dataset, and the other is the OCC annotation case dataset. The annotated datasets are divided into training set and test set in a ratio of 8:2. The divided datasets include 9,001 training samples and 2,250 test samples. The training set is used to train the emotion classification model, and the test set is used to verify the training effect of the model.

3.4.2 Model hyperparameter settings

In machine learning and deep learning, training a network model includes learning model parameters and determining hyperparameters. This method is used to train the generated model, so that the generated model parameters can be updated in a timely manner based on the training results(Li &Wang, 2021). Taking a simple neural network as an example $f(x) = w_1x + w_2x + b,$, where w_1, w_2 and b are all model parameters, they are initialized before training and updated continuously during training. Unlike the method of initializing model parameters, hyperparameter settings cannot be changed after being set. Common hyperparameters include learning rate, model depth, and model iteration times, etc. The setting of these parameters will greatly affect the learning efficiency of the model (Devlin et al., 2018).

The number of model iterations refers to the number of training iterations. If the number of model iterations is too large and the model continues to train before convergence, it will waste machine resources and time. If it is too small, the model may stop training before convergence, and the generalization effect of the model will be poor. In the case of an unspecified number of iterations, the model can be set to stop training when it does not update or decreases to the desired level within the required number of iterations. Setting the learning rate too high will cause the model to oscillate back and forth at a certain point and fail to converge, while setting it too low will cause the model to converge too slowly. If the optimal learning rate is not clear, the learning rate can be adjusted during the training process. A higher learning rate can be set for the early epochs, and a lower learning rate can be set for the later epochs. The setting of batch sample size is also very important. If the batch sample size is set too large, it will occupy too much memory and cause memory overflow. If it is set too small, it will slow down the training speed. Therefore, the specific batch data quantity needs to be determined based on the characteristics of the dataset and the hardware conditions (You et al., 2019). The setting of model hyperparameters used in this chapter is shown in Table 1.

Argument	Value
Validation split	0.2
Hidden size	768
Learning rate	5e-5
Attention head	12
Max position embeddings	512
Batch size	128
Epochs	10
Dropout rate	0.1
Loss	Categorical cross entropy
Optimizer	Adam

Table 1. BERT model part parameter Settings

Validation split represents the proportion of the validation set, hidden size represents the dimension of the hidden layer, learning rate represents the learning rate, attention head represents the number of attention heads. Max position embeddings represents the sequence length, set as 512 for BERT-base to support a maximum sentence length of 512, batch size represents the number of samples in each batch, epochs represents the number of iterations. Dropout rate represents the dropout rate, which is used in neural network training to temporarily remove neural network units from the network with a certain probability to prevent overfitting. Finally, the loss function is selected as the cross-entropy loss function, and the optimizer is chosen as "adam".

3.4.3 Experimental evaluation indicators

A trained sentiment classification model requires an evaluation indicator to assess its results. Through the evaluation indicator, the generalization ability of the model can be understood, and different models can be evaluated using the same indicator to determine which model has stronger generalization ability. The model parameters can be adjusted step by step through the evaluation indicator to optimize the model. In the field of deep learning, common evaluation indicators include confusion matrix, precision, accuracy, recall, and F1 value (Boursalie et al., 2021). The evaluation indicators used in this study are accuracy and F1 value, with accuracy as the main indicator.

Before describing the indicators such as accuracy, recall, and F1 value in detail, it is useful to briefly introduce some concepts and provide examples using the data used in the paper. Taking an example where a comment text belongs to the "positive sentiment" category, TP (True Positive) indicates that the comment text is correctly classified into the "positive sentiment" category; FP (False Positive) indicates that a comment text from a different category is incorrectly classified as "positive sentiment"; FN (False Negative) indicates that the comment text belongs to the "positive sentiment" category, but is classified into another category; TN (True Negative) indicates that the comment text is correctly classified as not belonging to any category other than "positive sentiment". P (Positive) and N (Negative) represent the actual classification made by the model, and T (True) and F (False) evaluate the correctness of the model's classification (Devlin et al., 2018). The relationship between TP, FP, FN, and TN is shown in Table 2.

Table 2. TP、FP、FN and TN

	Relevant	Non-relevant
Retrieved	TP	FP
Not retrieved	FN	TN

Accuracy rate refers to the proportion of the correct number of samples in the total number of samples, and the calculation formula is as follow.

$$Accuracu = \frac{TP + TN}{TP + FP + FN + TN}$$

Accuracy rate is also known as precision rate. For model prediction results, accuracy rate represents how many samples in the predicted positive samples are indeed positive. The ratio between the number of correctly classified positive samples and the number of classified positive samples is the accuracy rate. The calculation formula is shown as follows.

$$\Pr ecision = \frac{TP}{TP + FP}$$

The recall rate represents the proportion of the number of samples classified as positive samples to the number of all positive samples, which can also be understood as the number of samples in all positive samples predicted to be positive samples, and the calculation formula is as follows.

$$\operatorname{Re} call = \frac{TP}{TP + FN}$$

F1 value, also known as F1-SCORE, is based on the harmonic average of accuracy rate and recall rate, and the specific calculation formula is as follows (Pedregosa et al., 2011).

$$\frac{1}{F_1} = \frac{1}{2}\left(\frac{1}{P} + \frac{1}{R}\right)$$

$$F_1 = \frac{2 \times P \times R}{P + R} = \frac{2TP}{2TP + FP + FN}$$

3.4.4 Comparison experiment with benchmark model

In this chapter, the previous sections mainly elaborate on the theoretical methods, construction, and evaluation metrics of the OCC-BERT sentiment classification model. Based on this, this section conducts experiments to compare the performance of traditional machine learning algorithms SVM and LSTM with the sentiment classification model studied in this paper, using evaluation metrics.

This chapter has presented research stages of the OCC sentiment theory and the BERT algorithm. For this thesis, scientific validation of the OCC sentiment model and the BERT algorithm were conducted through experiments. The design idea of the comparative experiment was to divide the sudden event network public opinion text dataset into two categories: one category uses the OCC sentiment rules for sentiment annotation on the training set, and the other category does not use the OCC sentiment rules for sentiment annotation on the training set. The two training sets are used as the training sets for the comparative experiments of the BERT, SVM, and LSTM models, and the experimental results are compared and analyzed.

The process of SVM and LSTM classification algorithms is as follows: first, on the preprocessed dataset, use the jieba word segmentation tool to segment the burst event network public opinion text and filter out stop words; second, for these two benchmark models, use the Word2Vec word embedding method to construct word vectors, thus transforming the text data into a form that can be recognized by computers; finally, input the processed word vector data into the SVM and LSTM classifiers to obtain the sentiment classification results of the dataset. The experimental results obtained by BERT and these two benchmark models on the two categories of datasets are shown in Table 3.

	OCC annotated data set		Non-occ annotated data sets	
	Accuracy rate	F1	Accuracy rate	F1
BERT	0.8931	0.8931	0.8856	0.8852
SVM	0.7627	0.7582	0.7511	0.7453
LSTM	0.8242	0.8191	0.8149	0.8097

Table 3. Comparison of experimental results

In Table 3., by comparing the accuracy rates and F1 values of the three models, it was found that the BERT algorithm is significantly higher than the SVM and LSTM algorithms in terms of both accuracy rate and F1 value. This indicates that, from the perspective of comparing the three algorithms, BERT is more suitable for sentiment classification of burst event network public opinion and has better performance. At the same time, by comparing the OCC annotated dataset and the non-OCC annotated dataset, it can be found that after introducing the OCC sentiment model for sentiment annotation on the training set, the accuracy rates and F1 values of the three model algorithms have all improved. The accuracy rate of BERT has increased by 0.75 percentage points, SVM has increased by 1.16 percentage points, and LSTM has increased by 0.93 percentage points. Overall, the OCC dataset-based BERT sentiment classification algorithm proposed in this paper has the highest accuracy rate, demonstrating the superiority of the OCC-BERT model in the field of sentiment classification of burst event network public opinion.

3.4.5 Summary

This chapter first introduced the construction ideas of the burst event network public opinion sentiment classification model. Then, it provided an overview of the OCC sentiment model, simplified it, and presented designs burst event network public opinion sentiment rules. Next, it elaborated on the basic structure, pre-training, and fine-tuning of the BERT model, and proposed the combination of OCC and BERT algorithm to construct the burst event network public opinion sentiment classification model. Finally, through experiments, it verified that the BERT-based sentiment classification model proposed in this paper is superior to SVM, LSTM, and other algorithm models, and that the sentiment annotation of the comment dataset using the OCC sentiment model is accurate and reasonable, indicating that the OCC-BERT model has significant advantages in the field of sentiment classification of burst event network public opinion.

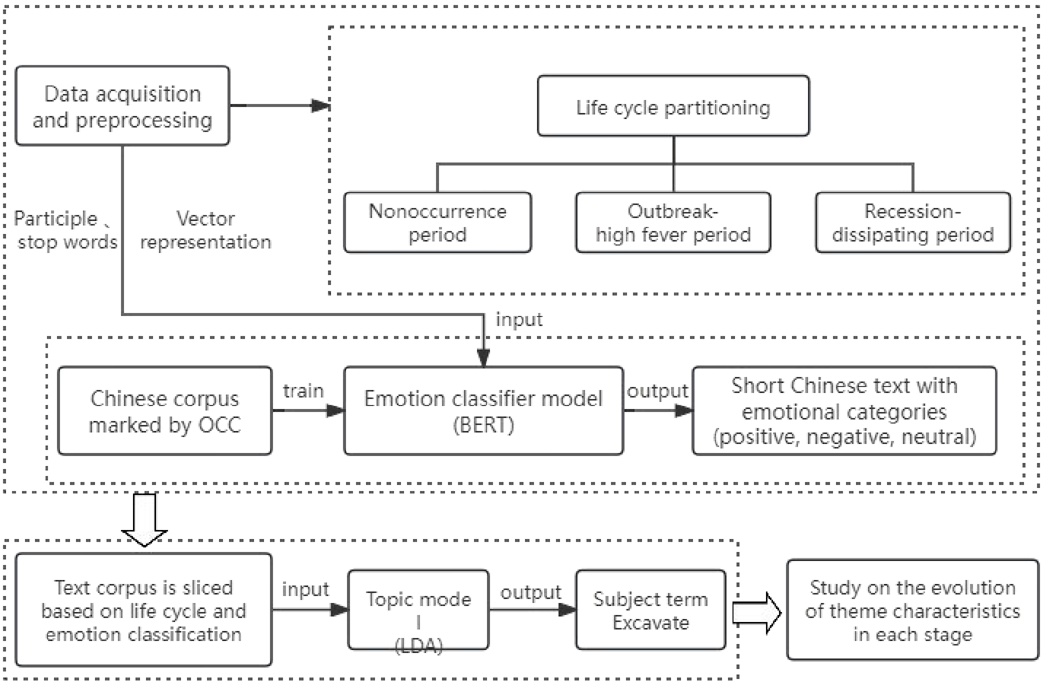
3.5 Topic analysis of sports events network public opinion based on LDA

This section presents a sports event network public opinion sentiment classification model based on the previous discussion. It establishes a sports event network public opinion topic analysis model based on LDA. The material integrates characteristics of public opinion stage division, sentiment classification results, and topic mining results, providing model and data support for the empirical analysis in the next chapter in order to further analyze specific cases of sudden event network public opinion.

3.5.1 Model Construction Idea

The main contents of the event network public opinion topic analysis based on LDA (Latent Dirichlet Allocation) constructed in this paper include four parts: public opinion lifecycle division, sentiment classification division, LDA topic mining modeling, and the study of topic sentiment evolution at each stage. First, based on the lifecycle theory, the division rules for dividing different stages of public opinion propagation cycle are constructed, so that subsequent topic mining can be conducted by segmenting the corpus based on time for evolutionary analysis. Second, the OCC-BERT sentiment classification model constructed in the previous chapter provides the sentiment classification results of public opinion texts. Therefore, the sentiment classification results and stage division are used to segment the LDA topic mining dataset as the input data for the next step of topic mining modeling. Third, based on the LDA model, topic mining is conducted by organizing the dataset according to the sentiment classification results and public opinion cycle division, and model training and topic mining are performed on different datasets. Fourth, based on the topic mining results of the previous step, combined with the collected corpus context, the feature words are summarized and summarized, and the topic feature evolution of each stage is studied. According to the above steps, this study utilized the framework of the sports event network public opinion topic analysis model based on LDA as shown in Figure 6.

Figure 6. Sports events network public opinion theme analysis frame



3.5.2 Division of Public Opinion Lifecycle

There are significant differences between the online public opinion of emergencies and ordinary online public opinion(Zhang, 2012). Existing models of online public opinion lifecycle mainly include three stages, four stages, or multiple stages. However, emergencies have characteristics such as suddenness and sensitivity, which lead to the differentiation of the evolution process of online public opinion of emergencies from ordinary online public opinion. After the event occurs, a massive amount of related information quickly spreads and ferments on online platforms, forming the online public opinion of the emergency in a short period of time. Moreover, the social attention caused by such events is much higher than that of ordinary hot events, and the event fermentation speed is fast, usually without an incubation period. Therefore, based on the characteristics of fast outbreak and evolution speed of online public opinion of emergencies, this article divides the lifecycle stages of online public opinion of emergencies in a broad sense and proposes division rules and characteristics (Zhang et al., 2020):

(1) Initial outbreak period. After the emergency occurs, the number of Weibo posts and comments related to public opinion rapidly increases from zero, and the search index shows an explosive rise. However, the language is relatively scarce and single, and there are fewer types of topics. The continuous emergence of new followers and topics indicates that the dissemination situation is in the initial outbreak period(Zhao, 2018).

Considering the sudden nature of the event and the wide-ranging discussions and attention it immediately triggers among netizens, the early stage of online public opinion of emergencies is usually much shorter than the early stage of ordinary online public opinion and has no incubation period. Also, because the event information is not comprehensive, netizens have limited access to information resources related to the event. Therefore, the focus of netizens' attention is basically consistent. In addition, at this stage, the scope of public opinion dissemination also expands from small group organizations to external networks, as manifested in the increase in the heat of online public opinion (Yu & Ni, 2024).

(2) Peak period of outbreak. The information increment of online public opinion of emergencies reaches its maximum, and the number of posts and comments rapidly reaches its peak. Netizens' attention is at its highest, and the heat of online public opinion is at its highest. The number of public opinion topics increases dramatically, and netizens' discussion remains highly active, with new messages and comments continuously appearing, indicating that the online public opinion of the emergency has entered the peak period (Zhou et al., 2020).

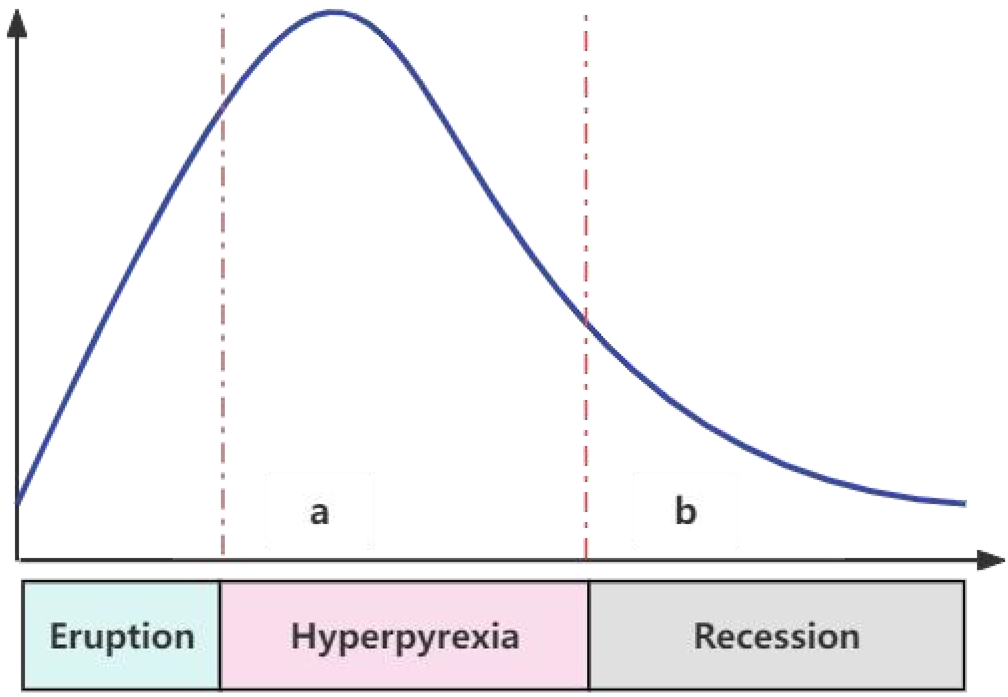
During the high-heat stage of public opinion, with the continuous exposure of information related to the emergency, netizens' attention to the event reaches its climax, and their communication and interaction with each other become more frequent, resulting in the rapid spread of public opinion in a viral manner. In this process, due to the intense discussions and excited emotions among netizens, a variety of opinions and claims collide, and similar opinions and claims gradually converge, resulting in a clustering effect. At the same time, the participation of mainstream media and opinion leaders makes it easy for netizens' attitudes to become one-sided, gradually leading to group polarization. Therefore, at this stage, relevant institutions should focus on studying the evolution of public opinion related to emergencies, grasp the changing trends of online public opinion, and timely identify potential crises in public opinion dissemination (Barnshaw et al., 2020).

(3) Decline and extinction period. When the amount of public opinion information rapidly decreases, both official and public postings and comments gradually decrease, and the number of new discussion points almost stops increasing, indicating a significant decline in the dissemination effect of public opinion and entering the decline stage of the online public opinion of the emergency (Qiu et al., 2019).

During the decline stage of public opinion, after a period of fermentation, the cause of the event and its disposal result have been revealed, and the focal issues of concern to the public have been properly resolved. As a result, news reports and online comments related to the event gradually decrease, and netizens' attention to the event quickly decreases. At the same time, netizens' emotions are alleviated to a certain extent, gradually moving towards rationalization. Although the online public opinion related to the event has gradually started to decline at this stage, it is still necessary to remain vigilant against derivative public opinions and prevent the rebound of online public opinion in emergencies(Ren & Li, 2023) .

When the amount of public opinion information almost stops increasing, the public clearly enters a period of slack in expressing attitudes and opinions, and public opinion approaches extinction, indicating the final stage of the online public opinion of the emergency (Hu, 2019).

Figure 7. Sports events network public opinion life cycle communication model



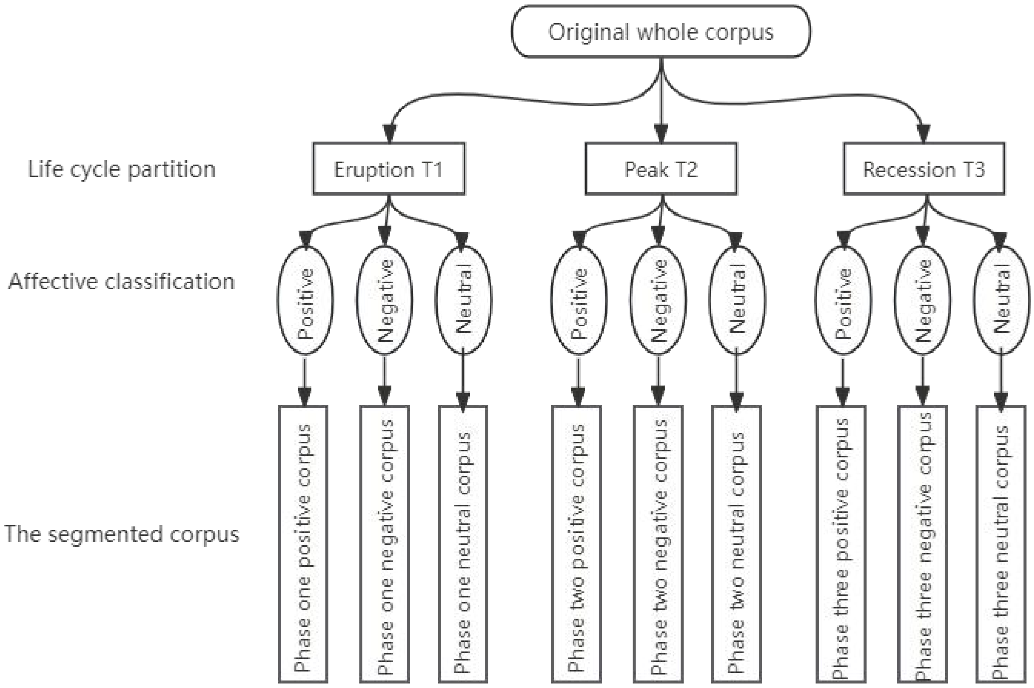
The figure 7. shows the overall performance of online public opinion during the lifecycle of sports events. The lifecycle diagram of public opinion on sports events reveals the law of dissemination. It uses time T as the horizontal axis and the amount of public opinion information as the vertical axis. The dashed lines a and b represent the boundaries between the outbreak stage, the high fever stage, and the decline stage. This paper establishes a universal lifecycle model, which can describe the lifecycle of public opinion dissemination on sports events under general conditions and make macro judgments on the development law of public opinion. Moreover, when dividing the lifecycle, there is no fixed pattern for the thinking of each scholar. Generally, each researcher can divide the lifecycle based on the specific communication situation of the event at that time (Gao, 2009). The case analysis section of this thesis will divide public opinion into these three stages using the example of skateboarding entering the Olympics. In addition, there have been some discussions about skateboarding before it officially entered the Olympics, so this paper will discuss the stages of the event two years before and after it occurred.

3.5.3 Emotion Classification Division

There are many uncertain factors in the development process of online public opinion on emergencies. Simply mining the themes directly from the public opinion texts can no longer meet the practical needs of current research on the evolution of public opinion. It is urgent to combine the emotions of the subject, the lifecycle of public opinion on emergencies, and theme mining in order to better grasp the focus of public attention and emotional evolution of netizens at different stages, take effective measures during critical time periods, guide public emotions and focus in the correct direction, so as to avoid the convergence of negative emotions and the emergence of secondary harmful events (Zhang et al., 2020).

Firstly, the lifecycle of online public opinion on emergencies is divided. Secondly, the preprocessed public opinion data is saved according to the time series. Then, the emotion classification model constructed before is used to classify the dataset by emotion and obtain the classification results sorted by time. Finally, the emotional categories of each stage within the lifecycle are counted, and the data is sliced according to the stages of public opinion and emotion categories, which can serve as the input for LDA theme mining. The division of emotional corpus in each stage is shown in figure 8.

Figure 8. Emotion corpus division in each stage



3.6 LDA Theme Mining

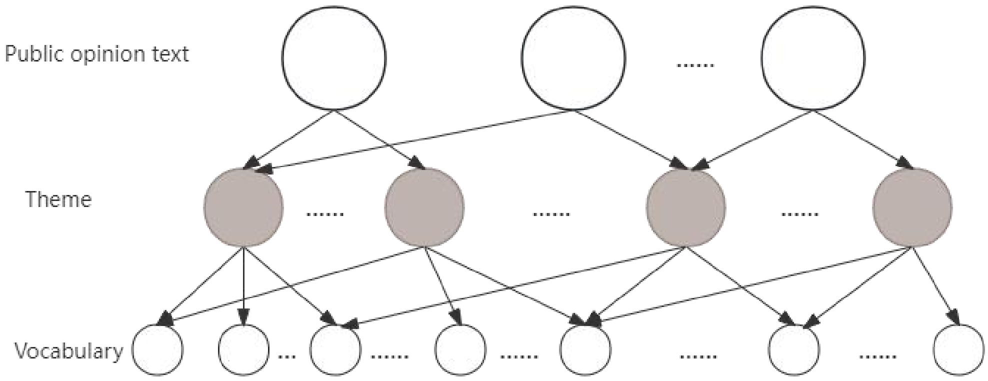
Throughout the evolution process of online public opinion on emergencies, there are often different discussion topics. The expression of emotions by netizens towards different topics also directly affects the speed and trend of the dissemination of emergencies. Based on this, this paper studies the evolution of themes in online public opinion on emergencies, analyzes the factors leading to the evolution of themes, and controls the development trend of online public opinion. In this chapter, I will use the Latent Dirichlet Allocation (LDA) model to mine the themes of online public opinion on emergencies (Cai et al., 2021).

3.6.1 LDA algorithm overview

(1) Introduction to LDA Structure. In 2003, Blei et al. proposed a more generalized text theme model called Latent Dirichlet Allocation (LDA). The essence of this model is a three-layer Bayesian probability model. In this paper, the three layers of the three-layer Bayesian theme model are the public opinion texts on emergencies, the themes, and the vocabulary. Assuming that the public opinion texts are composed of

multiple latent themes in certain proportions, all the vocabulary in the public opinion texts are formed into themes according to certain rules (Hu et al., 2023). The topological structure diagram is shown in figure 9.

Figure 9. Three-layer topology structure



(2)Algorithm Introduction.The LDA algorithm assumes that there are k independent themes in the public opinion text set D, and each theme follows a multinomial distribution in the vocabulary. Each public opinion text is a random mixture of k themes, and each public opinion text also follows a multinomial distribution in the k themes. Both the distribution of themes in public opinion texts and the distribution of vocabulary are Dirichlet prior distributions. The generation formula of LDA is as follows:

$$p(\vec{w},\vec{z}|\vec{\alpha},\vec{\beta}) = p(\vec{w}|\vec{z},\vec{\beta})p(\vec{z}|\vec{\alpha}) = \prod_{k=1}^{K}\frac{\Delta(\overrightarrow{\varphi_k}+\vec{\beta})}{\Delta(\vec{\beta})}\prod_{m=1}^{M}\frac{\Delta(\overrightarrow{\theta_m}+\vec{\alpha})}{\vec{\alpha}}$$

The text describes the topic generation process of sentiment analysis texts using the LDA algorithm as follows:

① Select a sentiment text d_i from the sentiment text set D according to the prior probability $p(d_i)$.

② Sample the Dirichlet distribution α to generate the topic distribution θ_i of sentiment text d_i.

③ Sample the topic multinomial distribution θ_i to generate the topic $z_{i,j}$ of the jth vocabulary in sentiment text d_i.

④ Sample the Dirichlet distribution β to generate the vocabulary distribution $z_{i,j}$ corresponding to the topic $\varphi_{zi,j}$.

⑤ Sample the vocabulary $\omega_{zi,j}$ from the multinomial distribution $\varphi_{zi,j}$ of the vocabulary.

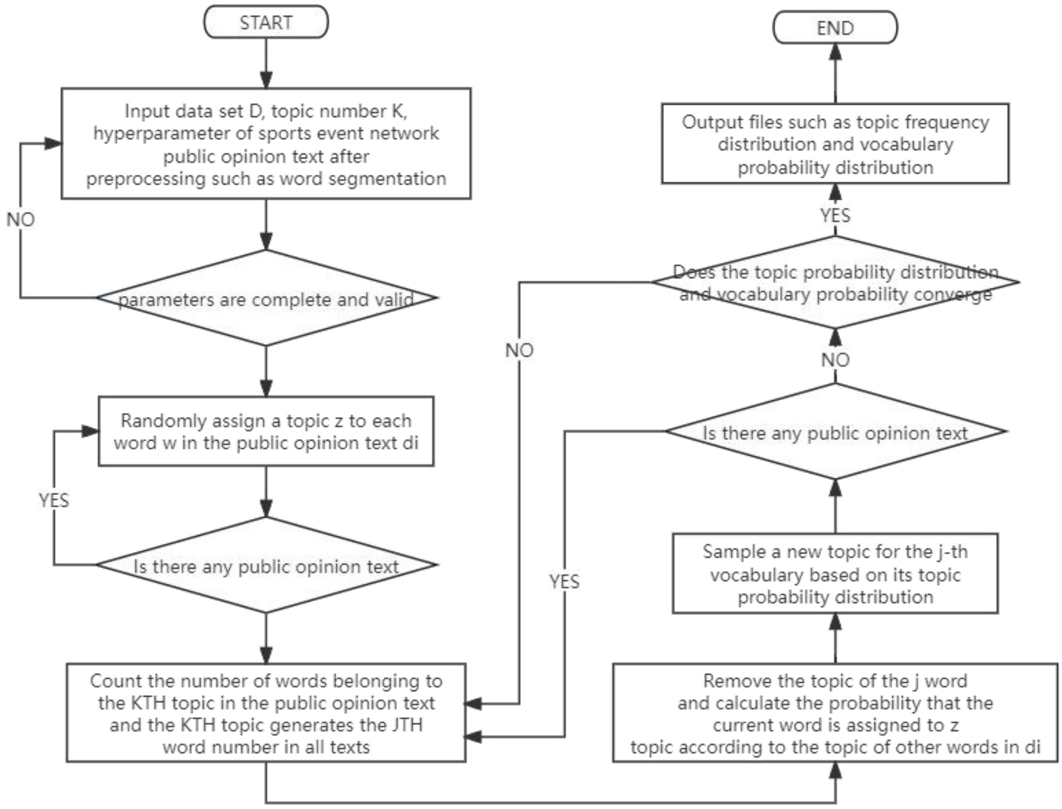
(3)The advantages of the algorithm are as follows:

LDA is based on the probabilistic latent semantic analysis model (PLSI) and adds two Dirichlet measures, which consider topic allocation and vocabulary allocation separately. This method not only efficiently identifies topics but also accurately characterizes the associations between topics and words. It can effectively reduce the dimensionality of topic features, thereby further improving the efficiency of network sentiment analysis. This method is an unsupervised machine learning method that can effectively solve the problem of instability in vector space models and overfitting of model parameters with the increase of the number of words. It has good model generalization ability. Based on these advantages, the LDA algorithm has a good prospect for related research on topic mining and has become one of the most widely used algorithms (Kherwa & Bansal, 2019). Therefore, this paper will use the LDA algorithm to mine topics in network sentiment analysis of sudden events.

3.6.2 Basic flow of LDA algorithm

The flow chart of topic mining for online public opinion of emergencies based on LDA algorithm is shown in Figure 10 (Xu et al., 2018).

Figure 10. LDA algorithm flowchart



In this study, the input part of the LDA algorithm mainly includes the dataset D of network sentiment texts of sudden events, the set parameters, and a reasonable number of topics K. The output part of the LDA algorithm mainly includes the assigned topic numbers of each vocabulary in each sentiment text, the probability distribution of vocabularies under each topic, the topic probability distribution of each sentiment text, and the high-frequency keywords under all topics. The operation steps of LDA are as follows:

① At the initial stage, input the preprocessed dataset of network sentiment texts of sudden events, and judge whether the parameters are complete and legal. If yes, proceed to the next step; if not, go back to the previous step to reset the parameters.

② Randomly assign a topic z to each word w in the network sentiment text id of sudden events until all words in the network sentiment texts of sudden events are assigned topics.

③Calculate the total number of keywords distributed in the kth topic in the network sentiment text id of sudden events, as well as the total number of the jth keyword generated by the kth topic in all network sentiment texts of sudden events.

④Filter out the topic of the jth word and calculate the possibility of the current word with topic z based on the topics of other words in the network sentiment text id of sudden events.

⑤Sample a new topic for this word based on the topic probability distribution of the jth word. Stop this process when all words in the network sentiment texts of sudden events have completed sampling for new topics.

⑥Obtain the updated values of the number of vocabularies corresponding to each topic and the probability values corresponding to each topic in the network sentiment text id of sudden events through multiple iterations and finally obtain the convergent result.

3.6.3 Parameter estimation

Parameter estimation is one of the key calculations in LDA topic modeling. In LDA, estimation of these two unknown parameters can be done using variational inference or Gibbs sampling. The former method, variational inference, follows the idea of maximum a posteriori estimation (treating the unknown parameters as fixed values), while the latter method, Gibbs sampling, follows the idea of Bayesian estimation. Bayesian estimation is an extension of MAP estimation but has a fundamental difference in that it considers the estimated parameters as random variables following a certain prior distribution. In comparison, Gibbs sampling is the most commonly used method in LDA topic modeling, and it has advantages such as fast computation speed and high estimation accuracy(Qiao & Zhang, 2000). Therefore, this study will also apply Gibbs sampling, and its calculation formula is as follows:

$$p(z_i = k|\overrightarrow{z_{-i}}, \vec{w}) \propto \frac{n_{m,-i}^{(k)} + \alpha_k}{\sum_{k=1}^{K} (n_{m,-i}^{(k)} + \alpha_k)} \cdot \frac{n_{k,-i}^{(t)} + \beta_t}{\sum_{t=1}^{V} (n_{k,-i}^{(t)} + \beta_t)}$$

3.6.4 Study on the evolution of subject characteristics in each stage

According to the division of public opinion stages and the construction of the emotion classification model, this paper uses the LDA algorithm to construct a research model for the topic characteristics evolution of public opinion on sudden events in different stages of the public opinion network. The analysis is mainly conducted from three aspects: the lifecycle of public opinion, text emotion, and topic mining (Li, 2017).

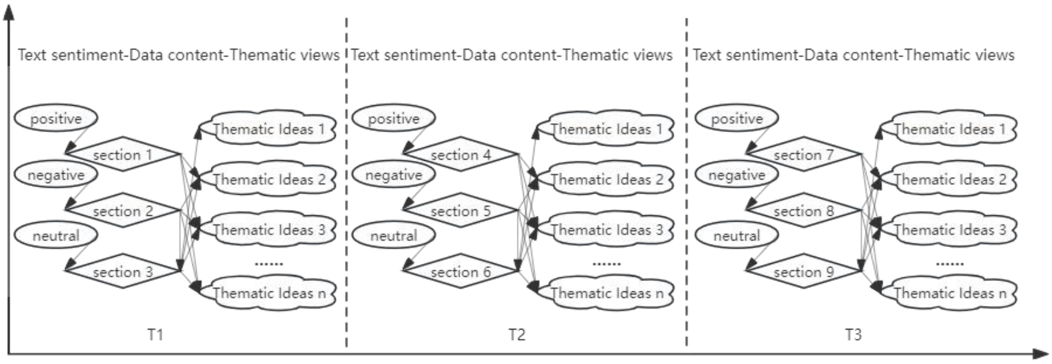
(1) Based on the publication of public opinion information and Baidu index statistics, this study divides the public opinion into stages with a time granularity of 1 day and conducts subsequent research on the first three stages with high information volume.

(2) Netizens' emotional attitudes towards a sudden event are reflected in the comments they post on online platforms.

(3) Based on the division of public opinion lifecycle, identification of emotional attitudes, and LDA topic mining of sudden event topics, the mining results of the topics are analyzed to summarize the perspectives on the topics.

From the above three aspects, public opinion data also has some characteristics: ① From the perspective of lifecycle, the themes held by public opinion events usually change in the short term, that is, in the same stage, there will be multiple perspectives on the topics. ②From the perspective of emotion, different emotions may contain different content at different periods, and different content may contain different perspectives on the topics. ③ From the perspective of topic perspectives, the relationship between data content and topic perspectives can be one-to-many or many-to-one. Therefore, this study establishes a research model for the evolution of topic characteristics in sudden event network public opinion by integrating the characteristics of public opinion data, including public opinion lifecycle, text emotion, and topic mining. The research model is shown in Figure 11. This study attempts to construct a research model for topic characteristics evolution, which can reflect the trends in the evolution of public opinion at different stages and under different emotions, and provide suggestions for the government to grasp and guide public opinion crises in a timely manner .

Figure 11. Research models of evolution of thematic features in different stages



3.6.5 Summary

This chapter first introduced the framework construction idea of the topic analysis of sudden event network public opinion based on LDA, and described the specific implementation steps. Secondly, based on the lifecycle theory, the division rules of different stages of public opinion propagation cycle were constructed. Thirdly, based on the results of the OCC-BERT emotion classification model constructed in the previous chapter and the stage division, the LDA topic mining data set was divided into corpora segments, and then the algorithm and process of the LDA topic model are described, and the topic model was trained. Finally, the research model for the evolution of topic characteristics in different stages was constructed by combining the division of public opinion lifecycle, the constructed emotion classification model, and the LDA topic model.

CHAPTER IV

SYATISTICAL ANALYSIS

This chapter uses "skateboard joining the Olympic Games" as a specific case study, and makes an empirical analysis of the OCC-BERT emotion classification model and LDA topic mining proposed above. Based on data acquisition and processing and evolutionary cycle division, emotional evolution analysis, topic mining analysis and theme feature evolution analysis were carried out. To provide relevant departments with effective evolutionary analysis of online public opinion for public opinion guidance suggestions.

4.1 Text vectorization

How to represent words as the first step in natural language processing tasks. In text classification, the first thing to do is to convert human-understandable natural language into machine-understandable data format, namely text vectorization. There are two commonly used methods for text vectorization: the bag-of-words model and the word embedding representation method(He, 2019). The bag-of-words model treats text as a combination of words, representing text as a bag of words. It does not consider the relationship between words and the order of words, meaning that each word is independent. The model assigns a weight to each word to represent the frequency of each word appearing in the text. Word embedding models are based on neural networks and represent word distributions. The model uses neural network techniques to construct context representations and the relationship between target words and contexts. It uses flexible neural networks to model complex contexts, resulting in word vectors that contain richer semantic information (Gallant, 1991).

Word2Vec is a common method for training word vectors, which simplifies the task of processing text content into vector operations in vector space. The similarity on the vector space represents the semantic similarity of the text. The model learns the semantic information of words from a large amount of text corpus, and the representation vectors of similar words are closer in the vector space, providing great assistance to downstream tasks (Deudon, 2018).

Word2Vec is a probabilistic language model based on neural networks, consisting of an input layer, a hidden layer, and an output layer. It is a three-layer shallow neural network. Word2Vec model has two structures: the Continuous Bag of Words (CBOW) model and the Skip-Gram model. Both models follow the aforementioned three-layer network structure. The CBOW model predicts the target word based on adjacent words. Its input is the word vectors of the context words adjacent to the target word, and the output is the word vector of the target word itself. The Skip-Gram model has the opposite input and output of the CBOW model. The input of the Skip-Gram model is the word vector of the target word itself, and the output is the word vectors of the adjacent words to the target word. The model predicts adjacent words based on the target word (Liu et al., 2015).

4.2 Overview of "Skateboarding added to Olympic Games"

The sport of skateboarding was officially included in the Olympic Games as four competitive events in the 2020 Tokyo Olympics. This marked the first appearance of skateboarding in the Olympics, signifying the widespread recognition of the sport's influence globally. This decision was formally approved during the International Olympic Committee meeting in August 2016. In addition to skateboarding, other sports such as surfing, climbing, baseball, and karate were also added. The inclusion of these sports has made the Olympics more diverse and attracted more participation from young people, while also reflecting the inclusivity and innovation of the Olympic Games (Wheaton & Thorpe, 2019). The discussion about extreme sports and the Olympics is ongoing, and even before skateboarding made its debut in the Tokyo Olympics, it had already generated considerable attention and discussion (Shi, 2015). Among the five new sports added to the Tokyo Olympics, skateboarding received more attention and generated more discussion than the others on platforms like Weibo. As a significant event related to the Olympics, the topic of "skateboarding joining the Olympics" generated a lot of online sentiment, from the initial discussions to its official debut in the Olympics, and will continue to be analyzed as a classic case study.

4.3 Data Acquisition and Processing

This study used web crawling techniques to collect and preprocess data on the topic of "skateboarding joining the Olympics" from the Weibo platform. The text data was then subjected to Chinese word segmentation and stop word filtering, and the OCC model was used to annotate the sentiment of the comment data, preparing the data for subsequent empirical analysis.

4.3.1 Data Acquisition

(1) Data Source. For the purpose of this study, data from Sina Weibo was selected as the experimental dataset. Considering that Sina Weibo is a popular social media platform, with a larger audience compared to portals like Netease News or Sina News. Sina Weibo users can post updates and comments anytime and anywhere using their mobile phones, providing greater flexibility. In addition, Sina Weibo has a wider range of data sources, as users can post original or reposted content based on their personal preferences, rather than being limited to commenting on official news like Netease. Furthermore, other open-source social tools like Tieba or blogs have certain limitations, as their content is more focused on specific areas or regions. Therefore, the data from Sina Weibo was chosen as the experimental dataset for this research. Weibo (similar to Instagram) is one of the most influential social media platforms in China, with a huge user base of 1 billion and a daily active user count of up to 300 million. Weibo users are known for quickly sharing their views on news and events, capturing real-time attitudes and reactions. The Weibo platform supports user interactions, including commenting and reposting, allowing for more user participation and reflecting a wide range of public opinions (Nichols, 2022).

In the rapidly developing era of the Internet, collecting and analyzing publicly available data plays an important role in industry development and technological advancement. In the field of natural language processing, sentiment analysis typically requires a large amount of data to uncover potential information from text, and using web crawling tools significantly reduces manual labor costs (Sun et al., 2016). Since there is no publicly available dataset of skateboarding comments online, the text was collected by crawling user comments on Weibo to construct a skateboarding evaluation dataset.

This study utilized web crawling to collect Weibo data. Web crawling involves three steps: obtaining the webpage, parsing the webpage, and storing the data. After determining the URL of the target webpage, an HTTP request is sent to the server using an HTTP library, waiting for a normal response from the server to retrieve the desired webpage content, resulting in a response that can be in the form of HTML, JSON, images, videos, etc. HTML data can be analyzed using regular expressions or third-party parsing libraries, JSON data can be parsed using the JSON module, and binary data such as images and videos can be saved to files for further processing. Valuable information is extracted from the parsed webpage and stored in a text file for subsequent use (Gao & Han, 2022).

(2) Data Overview. This study used web crawling to collect high-impact original popular Weibo posts and comment data using "skateboarding" and "Olympics" as keywords, from January 1st, 2019, to December 31st, 2022. The collected data was structured and saved in a spreadsheet, resulting in 2,351 original Weibo content data and 9,169 comment data, with only the Weibo posts with more than 1,000 comments being sampled at different time intervals. The main reason for selecting samples with more than 1,000 comments here is twofold. Firstly, Weibo posts with approximately 1000 comments maintain high popularity and are likely to appear at the top of search results. Secondly, some of the comments contain advertisements, and this number allows for filtering.

The collected Weibo dataset consists of publisher names, Weibo content, time, number of reposts, number of likes, and number of comments, while the comment dataset consists of user names, user profile page links, comment content, and comment time. Publisher and user names, user profile page links, Weibo and comment content are all strings, while comment and Weibo posting times are in date format, and the number of reposts, comments, and likes are numerical data.

Figure 12. Python collected part of the Weibo data graph

1	text
2	假装自己是滑板女孩很厉害的样子[偷笑]
3	宝马汽车授予香港雷星MINI品牌婴童自行车（10-16英寸）和滑板车等产品的全球独占性授权。星辉车模公司全资子公司香港雷星于8月17日同德国宝马汽车公司签订品牌授权合同！YETI非常适合
4	我的大宝贝回来啦，可把我给想坏啦我们俩见面抱了又抱不出所料他看到他的玩具区果然好开心呀还有姑姑给兑的滑板车，特别兴奋今晚我要抱着你睡觉[爱你][爱你][爱你]
5	一岁□ □ □ 天～今天在家忍着来M前的各种“伤痛”，坚持备课、听课[文明遛狗]猴宝贝昨晚玩滑板车摔了好几次，今天腿上全是伤，好难受好心疼啊！以后一定要更加小心地保护好我的
6	@手机用户1046628303 #迪卡侬儿童滑板Play120安全又好玩# [佩奇][乔治]
7	@空空的尾巴 @爱读书的人运气不会太差 @滑板少年的梦想
8	【心灵双约支教团2019团员风采 14】[心]王雨森 新闻部[心]爱好：美食 滑板 羽毛球[心]目标：做一个对得起自己的人[心]座右铭：天行健，君子以自强不息；地势坤，君子以厚德载物。[心]
9	KUSTOM STYLE2019 Spring & Summer CollectionFrom Yokohama Japan#改装车[超话]##滑板[超话]##纹身[超话]##kustomstyle##japan##wasterose##street#
10	最近每天换地方踏青，缝山针适合拍照，龙源湖适合玩滑板车，生态园适合放风筝，大沙河适合钓鱼还有野餐～最近经常被别人问有工作吗，其实一直有的，自己家开的辅导班，教小学奥数。说
11	33m+18[喜]6m+2 大上午出门玩的差事也只能落我肩上了，把拔每回都是睡睡睡，不到个大中午醒不过来。今天避高峰带大宝出门踏青，赏树山梨花，采摘草莓�没想到非周末人还是那么多。玩到
12	滑板车
13	新的一月，新的磨难，今天丢了最喜欢的耳机，滑滑板摔了一大跤，cnm
14	今天是个Cool boy神兽 烤鹅今天云真好看滑板还是不会过减速带 还把自己磕青了
15	今天玩了两个小时的滑板，感觉两个招式都有点意思了，哈哈，争取月中的时候可以保证出招次次成功！[哈哈]
16	#迪卡侬儿童滑板Play120安全又好玩#好想要一个送儿子@波波小鹿
17	//@社会滑板:带你看看碗池滑手的装备！ �� ����老鹰起飞！#社会滑板#
18	霹雳舞 进入了奥运会项目 哈哈[赞]还有滑板 我喜欢！
19	14M10D、喜欢坐在家里的各款小凳子上，不论高矮，都能爬上去坐好。甚至妈妈一坐她就哼哼叫，意思是坐了她的椅子。今天自己走去门口，指着鞋。拿到鞋坐地上自己开始揪袜子要换鞋。换好
20	滑板社退了自己也不滑了总不能占着位置不是回忆还是有的更多的是遗憾 没有混成一个圈子大一别扭的自己不愿意融入 又有点怵没一起刷过很远的地方 缝山针 景区什么的认识了有趣的人 但也
21	学会了滑滑板。
22	啊啊啊啊啊啊啊真的是想买滑板玩了
23	土豆滑板 手 机 壳！
24	遛狗➕滑板完美组合不用走路#livephoto[超话]##西安#
25	春天应该去海边滑滑板 涂鸦油画 随手摊开几本旧杂志开始看 和朋友去绿地里野餐 喝汽水吃蛋挞 穿着宽大的衬衫上街乱逛 听着hyukoh胡思乱想 短暂的去旅行 为学习而烦恼 但我没有喜欢的人
26	彩虹果冻Converse Chuck Taylor All Star 1970S 联名变色龙板鞋鞋身应用了台湾进口5D 变色材质，耐磨，在强光下面反光变色，实物更好看。Chunk 70s经典高帮休闲运动滑板板鞋“夜魔3M反
27	确实好厉害，我都不会玩滑板，� 但是话说回来，应该也是狗狗主人让它们学的，肯定也没少摔跤，狗狗为你们点赞！❤
28	#迪卡侬儿童滑板Play120安全又好玩#@贾小小小妞 新礼物
29	@Amy娜娜201701 #迪卡侬儿童滑板Play120安全又好玩#很好玩的
30	九点半孩子已经快睡着了，我猛地想起来今天没拍照片啊！每年的生日都有照片的呀！小时候乖乖的时候我们三个人都一起合影，现在他长大了，不爱拍照了，连哄带蒙的用相机给拍了几张，像

4.3.2 Data Processing

(1) Data cleaning. Firstly, the missing values were processed. Due to the large amount of data being crawled, a significant amount of data loss can occur. After investigation, it was found that some Weibo data did not have a posting time. In order to compensate for this, the earliest posting time of the corresponding comments was used to fill in the missing values. Since the posting time of the comments naturally comes after the posting time of the Weibo, the earlier the posting time of the comments, the closer it is to the Weibo time.

Secondly, the duplicate values were processed. Due to the fact that the Weibo text content and comment content are crawled on different pages, a large amount of duplicate text content is generated. This study used Excel functions to remove duplicate fields.

Thirdly, noise reduction processing. It was observed that the text data obtained from Weibo contains a large amount of irrelevant information. Therefore, this study used regular expressions to filter out the interfering information in the public opinion text data: ① removal of irrelevant symbols, numbers, and letters, such as "&", "-17", "f", etc.; ② deletion of URL hyperlinks in the text content, as these links have no semantic meaning in text comprehension and should be removed; ③ deletion of the content of "#hashtags", which are highly relevant terms mentioned by netizens during discussions but do not reflect their focus and are unrelated to the research; ④ "//@" indicates that this is a reposted content from others, which is unrelated to one's own viewpoint, so the content after "//@" is deleted; ⑤ simplification of traditional Chinese characters.

(2) Chinese word segmentation:

After the simplification of traditional Chinese characters, the removal of junk advertisements, and the data cleaning of invalid comments, the text data still cannot be directly used and needs to be segmented. Word segmentation refers to dividing a paragraph or sentence into several words that have practical semantic meaning. Currently, there are many applications of Chinese word segmentation tools (Wang, 2019). This study used the jieba word segmentation tool to segment the extracted Chinese data. Jieba is a Python open-source library specifically designed for Chinese word segmentation. There are three ways to segment words, namely precise mode, full mode (Wang & Liang, 2021), and search engine mode. The study used the precise mode by default to separate sentences as accurately as possible.

(3) Stop word filtering:

Stop words refer to words that are fundamentally useless or meaningless for the given purpose. They generally include conjunctions, adverbs, frequently occurring meaningless single Chinese characters, and terms like "you", "me", and "him". Although these words appear frequently, they have little or no use in extracting the main features of the text. If stop words are not filtered out, it will result in a large number of unrelated and ineffective messages, greatly affecting sentiment and topic analysis (Deshwal & Sharma, 2016). Therefore, after Chinese word segmentation, stop words need to be removed while retaining information related to sport events.

Currently, there are many versions of stop word lists, each with its own characteristics, including "Harbin Institute of Technology Stop Word List", "Baidu Stop Word List", and "Chinese Stop Word Library", etc. However, there is currently no clear stop word list that can be applied to various tools, which requires us to make different processing according to specific needs (Zhao, 2021). This study combined several stop word lists, including "Chinese Stop Word List", "Machine Intelligence Laboratory Stop Word List", "Baidu Stop Word List", and "Harbin Institute of Technology Stop Word List", to generate a stop word list containing 411 stop words. Some custom stop word lists are shown in Table 4.

Table 4. A partial display of the custom stop word list

?	<<	嘛	他的	到处	难怪
……	<Δ	什么	以前	顶多	偶尔
，	&	从	以致	顿时	其后
——	唉	从事	总的来看	非得	岂止
~	啦	从而	总而言之	赶快	恰恰
#	喏	他	恰恰相反	过于	恰巧

After word segmentation, there are still many meaningless words that appear frequently, such as "了" (le), "我" (I), and "的" (of). After stop word filtering, the segmented results show better effects. From Table 5, it can be seen that the processed text data hardly contains any invalid information.

Table 5. The result of removing the stop word segmentation

男孩子 四脚朝天 尖叫声 第一次
小时候 好想学 生命力
跟不上 荷尔蒙 小女生 小心翼翼 青少年
滑板车 安全带 消费者
自行车 第一次 快乐 轻松 刺激
腱鞘炎 真好玩 一点点
看不出来 第一次 救护车 平衡 舒适
自行车 激将法 刺激 学习

(4) Dataset sentiment labeling:

After data cleaning, a total of 53,234 comments remained. Among all the comment data, 4,251 pieces of data were randomly selected for OCC sentiment labeling, which served as the training dataset for the sentiment classification model. Based on the OCC sentiment rules, this study categorized the sentiment of public opinion comment text data into three major categories: positive sentiment, negative sentiment, and neutral sentiment.

In the case of the "Skateboarding Joins the Olympics" sports event, OCC rules were used to determine the sentiment of public opinion comments. For example, in the comments "Skateboarding is really fun" and "It's a good thing that skateboarding joins the Olympics," the event results align with public expectations, thus they are classified as positive sentiment. On the other hand, in the comment "Skateboarding needs to be trained properly, otherwise it is easy to get injured," the event result "easy to get injured" clearly does not meet public expectations, and "getting injured" is also against public behavior norms, therefore it is classified as negative sentiment. As for comments like "Is skateboarding a permanent event?" and "Are there any relevant championships for skateboarding?", they express concern and inquiry about the event, without involving the evaluation criteria of event results and behavioral norms. Therefore, they are classified as neutral sentiment. The OCC sentiment labeling of some comment data is shown in Table 5., with 1 representing positive sentiment, 0 representing neutral sentiment, and -1 representing negative sentiment.

Table 6. OCC rules label the emotional part of the data

Comment	Emotion classification
Skateboarding is really fun	1
Skateboarding needs to be trained properly, otherwise it is easy to get injured	-1
It's a good thing that skateboarding joins the Olympics	1
Is skateboarding a permanent event?	0
Are there any relevant championships for skateboarding?	0

4.4 Evolution cycle of network public opinion

The previous chapter, based on the life cycle theory, constructed a propagation model that divided sudden events into three stages: initial outbreak period, outbreak peak period, and decline phase. Here, combining various data such as public opinion information and Baidu search index, the specific dates of the three stages of public opinion are determined.

The Baidu search index data for the event of skateboarding joining the Olympics was used to create a time distribution chart, as shown in Figure 12. It can be seen that the attention to "skateboarding joining the Olympics" began to increase in March 2020, and the related search volume decreased daily afterwards. In July 2021, there was a small peak, which was because the skateboarding event of the Tokyo Olympics in 2020 took place in July 2021, bringing the topic back into the public eye and generating renewed interest.

Figure 12. Public opinion life cycle based on Baidu index

Based on the Baidu search index data for the Tianjin explosion on August 12th, it was found that the network public opinion life cycle of the "skateboarding joining the Olympics" event, divided based on the Baidu search index, was basically consistent with the trend based on the Baidu search index. Therefore, the network public opinion life cycle of the "skateboarding joining the Olympics" event is divided into: ① initial outbreak period, January 2019 - June 2021; ② peak period, June 2021 - July 2021; ③ decline phase, August 2021 - December 2022.

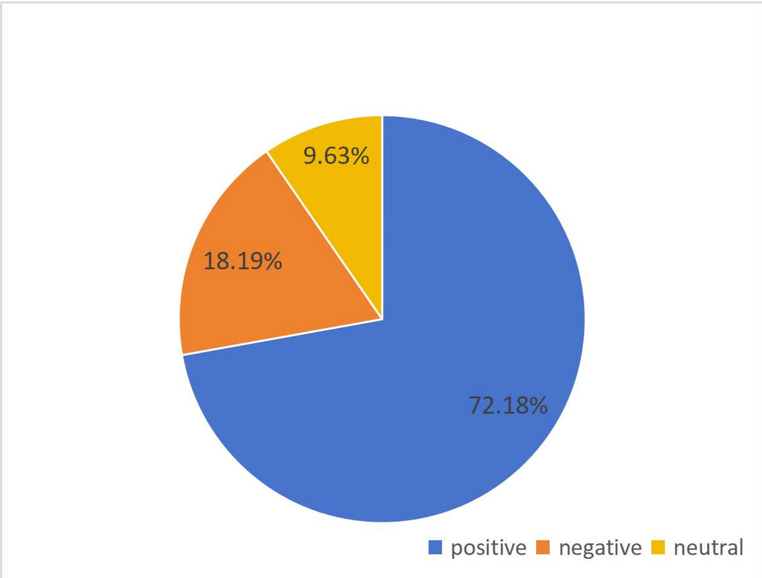
4.5 Sentiment evolution analysis of network public opinion

In this section the constructed OCC-BERT sentiment classification model is applied to predict the sentiment polarity of the text corpus of the "skateboarding joining the Olympics" event. After model training, predicted data with annotated sentiment polarity are obtained. Based on this, the sentiment evolution analysis of the "skateboarding joining the Olympics" event in network public opinion is conducted.

4.5.1 Overall sentiment evolution analysis

This study conducted a statistical analysis of the sentiment prediction results, categorizing sentiments into positive, negative, and neutral categories. As shown in Figure 13., it was found that overall, positive sentiment accounted for 72.18%, negative sentiment accounted for 18.19%, and neutral sentiment accounted for 9.63%. This indicates that the majority of netizens hold positive attitudes towards the "skateboarding joining the Olympics" event, the overall positive emotion prevails, and more detailed analysis of the specific emotion distribution throughout the life cycle is needed.

Figure 13. Overall affective category proportion



4.5.2 Stage sentiment evolution analysis

Analysis of Table 7. and Figure 14. reveals that in the initial outbreak stage, positive sentiment dominates, indicating that those who paid attention to the event of skateboarding joining the Olympics were mostly skateboard enthusiasts. In the peak stage, people's emotions gradually tend towards positive sentiment, and there are obvious differences in emotions. Positive sentiment and negative sentiment are almost equal, indicating a significant divergence. One reason may be that skateboarding, as an extreme sport, although it has some enthusiasts, many people have limited knowledge about skateboarding, and there are many doubts among netizens about skateboarding injuries, leading to a greater division of emotions among the public. In the third stage, the decline phase, there is not much change in neutral emotions, but the proportion of positive sentiment compared to negative sentiment increases. After skateboarding entered the public's field of vision, many popular science videos and articles appeared on the Internet, which further enhanced the public's understanding of skateboarding. As a result, negative emotions among netizens gradually decreased while positive emotions increased, and a positive sentiment began to emerge among netizens, leading to a more relaxed mindset. Despite the gradual decrease in the intensity of discussions among netizens in recent times, sports events are being held periodically, so there may still be a resurgence of public sentiment towards skateboarding during periods of decline.

Table 7. The proportion of emotion in each stage

Phase	Proportion		
	Positive	Negative	Neutral
T1	0.63	0.21	0.16
T2	0.88	0.09	0.03
T3	0.74	0.12	0.14

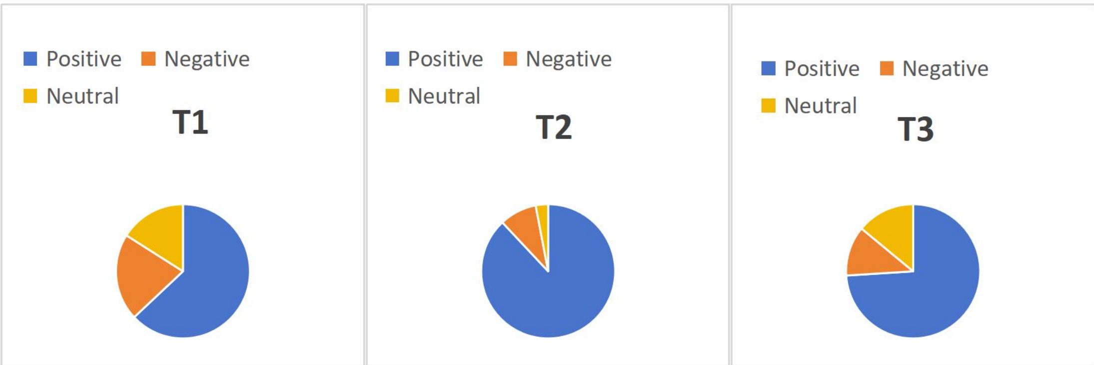


Figure 14. Emotional distribution at each stage

4.6 Network public opinion theme analysis

The previous section divided the life cycle stages based on Baidu index and Weibo data volume, and carried out text emotion recognition analysis based on OCC-BERT deep learning model.In this section, the corpus will be divided by life cycle and emotion. Based on the LDA topic mining method constructed in the previous chapter, the public opinion texts under different emotional categories at each stage of the public opinion of "skateboard project joining the Olympic Games" will be subject mined and their subject contents analyzed.

4.6.1 Initial outbreak period

The initial stage of the outbreak of online public opinion of this sports event was from January 2019 to June 2021. The topics under this stage are explored and summarized, as shown in the Table 8.

Table 8. The results of topic mining in the preliminary outbreak stage

Phase	Partial subject term	Summary of topic content
Initial outbreak period	P1：安全、设计、好玩、酷炫	P1：Properties of skateboard
	P2：妈妈、弟弟、晚上、微风、朋友	P2：Skateboarding social relations
	P3：开心、喜欢、有趣、刺激	P3：The embodiment of skateboarding experience

The initial outbreak stage T1 is known that skateboarding has joined the Olympic Games but has not made an appearance in the Olympic Games. Most participants in the sport of skateboarding are skateboarding enthusiasts. In long-term participation in skateboarding, most of them are positive emotions, among which negative emotions are mostly shown by people injured while skateboarding.Most of the public attention is on the attributes of skateboarding itself and the direct experience brought by participating in skateboarding events. Long-term participants mostly record the daily life of participating in skateboarding projects.The public opinion of neutral emotion is mostly because the network push paid attention to the skateboard project for the first time, and did not find emotion-related remarks on the basis of not understanding them.

4.6.2 Eruption period

The outbreak stage of online public opinion of this sports event was July to August 2021. The theme of this stage is explored and summarized, as shown in the Table 9.

Table 9. The results of topic mining in the outbreak stage

Phase	Partial subject term	Summary of topic content
Outbreak period	P1：联名、奥运会、曾文惠、奥运选手	P1：Skateboarding and the Olympics
	P2：刺激、有趣、尝试、舒爽	P2：Skateboard positive emotion
	P3：受伤、医院、保险、医药箱	P3：Negative emotions of skateboarding

The outbreak phase T2 occurred when the sport of skateboard officially appeared in the Tokyo Olympic Games and its competition stage. During the Olympic Games skateboarding events live broadcasts, many people who pay attention to the Olympic Games in China did not understand the rules and scoring of the sport. Live TV assisted them to understand, and generated heated Internet discussion.The positive discussion mainly focused on the experience after watched the live broadcast and the extreme sports experience after participating in the skateboarding project for the first time. The negative discussion was still the problem of sports injury. At the same time, the discussion about the Olympic Games skyrocketed and began to focus on Chinese skateboarders, and the direct connection between skateboarding and the Olympic Games increased.

4.6.3 Recessionary extinction period

The decline and extinction period of the network public opinion of this sports event is from September 2021 to December 2022. The topics in this period are explored and summarized, as shown in the Table 9.

Table 10. The results of topic mining in the recessionary extinction stage

Phase	Partial subject term	Summary of topic content
Recessionary extinction period	P1：设计、安全、美观、舒适性	P1：About the skateboarding project itself
	P2：哥哥、朋友、社区、傍晚	P2：Participate in skateboarding social relationships
	P3：医院、摔伤、滑雪、危险、医生	P3：Negative emotions of skateboarding

The decline and demise of T3 is the stage after the Olympic Games, netizens have a preliminary understanding of skateboarding because of the Olympic Games, the subsequent skateboarding related events have been carried out, enthusiasts continue to pay attention to, it may cause every time Chinese athletes participate in skateboarding competitions, it will cause public opinion hot spots on the network.

4.7 Analysis on the evolution characteristics of the emotional theme of network public opinion

Through the analysis of Weibo comments on different stages and emotions of the "Skateboarding in the Olympics" public opinion event, this study selected high-probability feature words under each emotional theme in the lifecycle for interpretation and thematic feature induction results.

From the initial outbreak stage, it can be seen that netizens mainly engaged in event discussions and express emotions. In the positive emotion corpus, words such as “frequently participating in skateboarding”, “making new friends through skateboarding”, and “skateboarding in leisure time” express a positive attitude towards the development of time. In the negative emotion corpus, words like “injury” and “hospital” represent the negative emotions of netizens during this period. In the neutral emotion corpus, it is understood that netizens receive information through news, watching videos, and other means, including the main information related to the event.

From the outbreak stage, it can be seen that netizens mainly engaged in event tracking and express emotions. In the positive emotion corpus, words such as “exciting skateboarding”, “young sports”, and “want to participate” added the willingness and behavior of wanting to participate in sports compared to the previous stage expressing good wishes. In the negative emotion corpus, words such as “first aid kit”, “falling”, and “injury”, in addition to expressing fear, also added a psychological state of worrying about getting injured during the participation process. In the neutral emotion corpus, there were discussions on the new sports added to the Olympics.

From the decline stage, it can be seen that netizens mainly engage in post-events arrangements and express emotions. In the positive emotion corpus, it was understood that the number of people who like skateboarding has increased and skateboarders have received attention. In the negative emotion corpus, it expressed understanding for the injuries caused by participating in skateboarding. In the neutral emotion corpus, it conveyed netizens' discussions on the addition of new events to the Olympics.

Based on the above, in the process of public opinion communication, the thematic views of the communication topics of positive emotion corpus, negative emotion corpus and neutral emotion corpus have their own characteristics emotional tendency and reasons. Meanwhile, in terms of the overall evolution of public opinion at each stage, netizens' concerns vary from "event discussion" to "event tracking" to "event arrangement", and each stage is accompanied by certain "emotion expression".But the emotions expressed change with each stage.It can be seen from the evolution of the characteristics of the public opinion communication cycle that the feature evolution excavated based on the research case is in line with the logic of the development of events.Therefore, it shows that the research model of the evolution characteristics of online public opinion on emergencies based on life cycle, emotion and topic data is an effective research tool.

4.8 Summary

This chapter used Weibo and Baidu Index as the data sources of online public opinion for sport events, took the online public opinion of skateboard's inclusion in the Olympic Games as case study, conducted a series of pre-processing of the collected data, and divided the life cycle of public opinion based on the life cycle and the amount of public opinion information.The overall and stage emotion evolution of OCC-BERT model was analyzed for the emotion recognition results of public opinion text data.At the same time, LDA model was used to explore and summarize the theme of skateboard joining the Olympic Games, analyzed the theme evolution characteristics of network public opinion, and finally put forward policy suggestions for empirical analysis.

CHAPTER V

DISCUSSION & SUGGESTIONS FOR FUTURE RESEARCH

5.1 discussion about skateboarding

In the context of rapid development in global socio-economics and globalization, the frequency of various events erupting continues to increase, bringing significant heat and discussion to public life and society. In today's rapidly developing internet era, sports, with its unique position and role, leads the world trend and is closely linked to various countries, national economies, and politics (Rubdy, 2001). Although the discussion about whether the Olympics should stay away from political factors has never stopped, in today's era of globalization, sports inevitably have multiple connections with various aspects of society (Gu, 2008). Therefore, monitoring online public opinion has become increasingly important. The Olympics continue to innovate by adding new events and making changes based on feedback and the development of the times. It is also hoped that the Olympics can adapt to the development of the times, attract young people and people of all ages to participate in sports, and enhance the Olympics' attention and influence.

This thesis takes sports event online public opinion as the specific research object and conducted research on the evolution of public opinion sentiment and topic features from the perspective of text sentiment analysis and topic mining. By using the lifecycle theory and dividing public opinion information into different stages, the OCC-BERT model was used to classify and extract sentiment from public opinion data(Liu & Gu, 2022), and it was applied in the case study to realize the analysis of the emotional evolution of online public opinion. By dividing the corpus into slices based on lifecycle and sentiment classification, the LDA method was used to identify and extract public opinion topics, and the results of topic mining that integrate sentiment and lifecycle were identified and summarized(Chen et al., 2015), accurately was deconstructing the topic features of sports event online public opinions and exploring the changes in netizens' emotions and focal points. Based on the research ideas mentioned above, a case study on the inclusion of skateboarding in the

Olympics was conducted for empirical analysis. Through the construction of the aforementioned models and empirical analysis, the main research conclusions obtained in this article are as follows:

(1) In the analysis of the emotional evolution of online public opinion regarding the inclusion of skateboarding in the Olympic Games, significant differences in the emotional tendencies of netizens were found at different time periods. During the initial stage of the event, most netizens were relatively rational and expressed positive emotions, such as anticipation for skateboarding to be officially included in the Olympic Games. The negative emotions expressed by netizens mainly centered around injured individuals. After the official debut of skateboarding at the Tokyo Olympics, the intensity of negative emotions expressed by netizens was significantly smaller compared to positive emotions. Additionally, regardless of the mainstream stage of positive emotions or negative emotions, the development process will produce contradictory emotions, positive and negative emotions always exist simultaneously. During times when negative public opinion increased, the group of skateboarding enthusiasts began posting publish educational content on social media platforms to popularize skateboarding in school and leisure time.

(2) In the analysis of the thematic features of online public opinion regarding the inclusion of skateboarding in the Olympic Games, changes in netizens' focal points were observed under different stages and emotions. By mining netizens' focal points, the reasons for the changes in netizens' emotions were reflected. With the subsequent development of the event and the passage of time, discussions, tracking, and arrangement of skateboarding relevant topics surrounding the event took the lead, and netizens' emotional expressions were present in every stage. At the same time, from the thematic expressions of netizens' negative emotions, it can be seen that the main source of negative public opinion is the potential sports injuries associated with skateboarding itself. Therefore, it can be inferred that the negative emotions in skateboarding make it necessary to consider many factors in the future development and promotion of skateboarding. Attention to skateboarding injuries can be extended to the teaching, training, and research of skateboarding. As one of the popular options

for parents and students to choose as a sport, skateboarding is one of the extreme sports with relatively low participation difficulty specific venue requirements. However, the training and teaching market is relatively chaotic(Wang, 2022), and the qualification of coaches varies. Despite the low venue requirements, there are few venues that are truly well-equipped and guaranteed, and the operation and management of venues also need to be planned for the long term. Finally, the athletes in the skateboarding project are relatively young, and the selection and training of future athletes not only require the learning of professional skills but also need to follow the growth rules of adolescents in terms of curriculum arrangement, and the learning of cultural classes and skateboard training should be carried out at the same time.

(3) The OCC-BERT sentiment classification model can accurately identify the sentiment tendencies of microblog texts in sports event online public opinion. The sentiment recognition model based on the BERT model constructed for sport event online public opinion is superior to the LSTM algorithm based on deep learning and the machine learning support vector machine (SVM) algorithm in terms of sentiment recognition effectiveness. In the process of sentiment annotation of online public opinion texts of the event, the application of OCC sentiment rules can improve the accuracy of the sentiment recognition model. The combination of the OCC sentiment rules and the BERT model can achieve a maximum accuracy of 0.8931. This indicates that the OCC model and the deep learning BERT model have certain scientific and advanced features in the sentiment classification research of burst event online public opinion. They provide a support basis for the analysis and decision-making of sports projects, and also offer a new solution for the research of sentiment classification in online public opinion.

5.2 Limitation

Under the limitation of time and energy, this study has limitation. It can be further improved in two aspects. Firstly, although this paper uses the advanced BERT model combined with OCC sentiment rules for sentiment classification of public opinion text content in the emotion recognition part, it neglects the word position in the text by using the topic model (LDA) based on machine learning algorithm for topic mining. Additionally, the results of the topic model (LDA) are greatly influenced by the length of the text, and for short texts, the frequency of word occurrence is usually low(Qiu, 2024). With the widespread application of deep learning in the field of natural language processing, topic extraction based on deep learning will have better extraction effect than the topic model (LDA), and can explore the association attributes between words to make the semantic information more abundant.

On the other hand, in the selection of public opinion events, only the topic "skateboarding entering the Olympics" is analyzed, and it is necessary to combine the public opinion events of four other sports entering the Olympics to further verify the sentiment evolution and topic feature analysis of the sports event network public opinion research methodology. Subsequent research could also analyze multiple public opinion events in order to optimize the research on sentiment evolution and topic feature analysis of public opinion as proposed in this paper, and meet the analysis needs of a large number of public opinion events in the sports event context.

5.3 Suggestion for future research

With the successful inclusion of skateboarding in the Olympics, its development in China has encountered new opportunities and challenges. As a vibrant and promising sport, skateboarding has a broad prospect for development in China. However, in order to realize this potential, a series of measures need to be taken to promote the long-term development of skateboarding in China. This article will propose suggestions regarding the current status, challenges, and future development directions of skateboarding in China.

1. Strengthen infrastructure construction. The development of skateboarding requires good infrastructure support, including skate parks, public skateboarding facilities, and related supporting facilities(Wang, 2023). Therefore, it is recommended that the government increase investment in skate park construction, provide more skate parks and skateboarding areas, and offer more practice and communication venues for

skateboarding enthusiasts. It is also important to enhance the management of skate parks and skateboarding areas, and equip them with medical measures.

2. Promote the popularization of skateboarding culture. Skateboarding is not only a sport but also a culture and way of life(Dawid, 2009). Therefore, it is recommended to strengthen the promotion and popularization of skateboarding culture through organizing skateboarding competitions, cultural activities, promotional events, and other means. This will help more people understand and accept skateboarding culture, thereby increasing the audience base for skateboarding.

3. Develop skateboarding education for youth. Youth are the main participants and driving force of skateboarding. It is suggested to strengthen training and education for youth skateboarding, promote skateboarding courses and activities in schools and communities, cultivate more outstanding young skateboarders, and provide them with development opportunities and support.

4. Enhance international exchanges and cooperation. As an international sport, skateboarding needs to be connected with the international community and strengthen international exchanges and cooperation. It is recommended to strengthen connections with international skateboarding organizations, actively participate in international skateboarding competitions and events, and enhance the influence and competitiveness of Chinese skateboarding on the international stage.

5. Cultivate excellent athletes and coaching teams. Talent identification and development are key to the development of skateboarding. It is suggested to strengthen the training and selection of skateboarders and coaches, improve their competitive and teaching levels, and provide more outstanding talents for the long-term development of skateboarding.

6. Enhance the construction of the industry chain. The development of skateboarding also requires a sound industry chain support. It is recommended to strengthen the development of related industries such as skateboarding equipment, clothing, and event organization, promote the healthy development of the skateboarding industry chain, to provide more support and guarantee for skateboarding.

7. Strengthen scientific research and popularization. Skateboarding is a complex sport, and the sustainable development of skateboarding requires continuous scientific research support. It is recommended to strengthen scientific research and popularization of skateboarding, explore the laws and characteristics of skateboarding, provide scientific training methods and guidance to promote the further development of skateboarding.

In summary, the development of skateboarding in China faces many challenges but also holds enormous potential for development. By strengthening infrastructure construction, promoting the popularization of skateboarding culture, developing skateboarding education for youth, enhancing international exchanges and cooperation, cultivating excellent athletes and coaching teams, enhancing the construction of the industry chain, and strengthening scientific research and popularization, skateboarding can be promoted for long-term development in China and achieve success on the international stage.

BIBLIOGRAPHY

Adam, C., Herzig, A., Longin, & D. (2009). A logical formalization of the occ theory of emotions. SYNTHESE -DORDRECHT-.

Barnshaw, J., Letukas, L., & Quarantelli, E. L. (2008). The characteristics of catastrophes and their social evolution: An exploratory analysis of implications for crisis policies and emergency management procedures.

Batuev, M., & Robinson, L. (2017). How skateboarding made it to the Olympics: an institutional perspective. International journal of sport management and marketing, 17(4-6), 381-402.

Beal, B. (2023). Searching for Relevance: The Olympic Games and Action Sports: A Review Essay of Belinda Wheaton and Holly Thorpe, Action Sports and the Olympic Games.

Beal, B., Atencio, M., Wright, E. M., & McClain, Z. (2017). Skateboarding, community and urban politics: shifting practices and challenges. International Journal of Sport Policy and Politics, 9(1), 11-23.

Boursalie, O., Samavi, R., & Doyle, T. E. (2021, February). Evaluation metrics for deep learning imputation models. In International Workshop on Health Intelligence (pp. 309-322). Cham: Springer International Publishing.

Cai, M., Luo, H., & Cui, Y. (2021). A study on the topic-sentiment evolution and diffusion in time series of public opinion derived from emergencies. Complexity, 2021, 1-23.

Chen, S., Yang, J., Yang, W., Wang, C., & Bärnighausen, T. (2020). COVID-19 control in China during mass population movements at New Year. The Lancet, 395(10226), 764-766.

China Internet Network Information Center. (2023). 51st Statistical Report on the Development of China's Internet. Retrieved from http://www.cnnic.cn/

Coalter, F., & Taylor, J. (2009). Large scale sports events: event impact framework, report to UK sport.

D'Orazio, D. (2021). Skateboarding's Olympic moment: The gendered contours of sportification. Journal of Sport and Social Issues, 45(5), 395-425.

De Choudhury, M., Counts, S., & Gamon, M. (2012). Not all moods are created equal! exploring human emotional states in social media. In Proceedings of the International AAAI Conference on Web and Social Media (Vol. 6, No. 1, pp. 66-73).

Deshwal, A., & Sharma, S. K. (2016, September). Twitter sentiment analysis using various classification algorithms. In 2016 5th international conference on reliability, infocom technologies and optimization (Trends and Future Directions)(ICRITO) (pp. 251-257). IEEE.

Deudon, M. (2018). Learning semantic similarity in a continuous space. Advances in neural information processing systems, 31.

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805.

Exploratory Study of International Sports Event Life Cycle. (2019). Journal of Sport and Applied Science, 3(1), 1–14. https://doi.org/10.13106/KJSAS.2019.VOL3.NO1.1.

Foster, G., O'Reilly, N., Shimizu, C., Khosla, N., & Murray, R. (2014). Determinants of regional sport network television ratings in MLB, NBA, and NHL. Journal of Sport Management, 28(3), 356-375.

Gallant, S. I. (1991). A practical approach for representing context and for performing word sense disambiguation using neural networks. Neural Computation, 3(3), 293-309.

Gao, P., & Han, H. (2022). Robust web data extraction based on weighted path-layer similarity. Journal of Computer Information Systems, 62(3), 536-546.

Garcia, B. (2008). One hundred years of cultural programming within the Olympic Games (1912–2012): origins, evolution and projections. International journal of cultural policy, 14(4), 361-376.

Gazeres, R. (2023). Challenging neoliberal sport: Skateboarding as a resilient cultural practice. Geography Compass, 17(1), e12671.

Gil-García, R., & Pons-Porrata, A. (2010). Dynamic hierarchical algorithms for document clustering. Pattern Recognition Letters, 31(6), 469-477.

Hu, C., Liang, Q., Luo, N., & Lu, S. (2023). Topic-Clustering Model with Temporal Distribution for Public Opinion Topic Analysis of Geospatial Social Media Data. ISPRS International Journal of Geo-Information, 12(7), 274.

Hu, Y. (2019). Text mining and data information analysis for network public opinion. Data Science Journal, 18, 7-7.

Kherwa, P., & Bansal, P. (2019). Topic modeling: a comprehensive review. EAI Endorsed transactions on scalable information systems, 7(24).

Kwon, K., & Oh, T. (2019). Exploratory study of international sports event life cycle. Journal of Sport and Applied Science, 3(1), 1-14.

Liu, P., Qiu, X., & Huang, X. (2015, July). Learning Context-Sensitive Word Embeddings with Neural Tensor Skip-Gram Model. In IJCAI (pp. 1284-1290).

López, C. Q., Vera, V. D. G., López, C. V., & Jaramillo, J. C. Á. (2020). Impact of sport on social cognition: An analysis based on structural equational models. Journal of Physical Education and Sport, 20(1), 31-36.

Malfas, M., Theodoraki, E., & Houlihan, B. (2004, September). Impacts of the Olympic Games as mega-events. In Proceedings of the Institution of Civil Engineers-Municipal Engineer (Vol. 157, No. 3, pp. 209-220). Thomas Telford Ltd.

Naud, A., & Usui, S. (2007, December). Exploration of a text collection and identification of topics by clustering. In International Conference on Intelligent Data Engineering and Automated Learning (pp. 115-124). Berlin, Heidelberg: Springer Berlin Heidelberg.

O'Connor, M., Mehta, S., & Milanaik, R. (2022). The Lack of Helmets in Olympic Skateboarding—Skating Into Dangerous Territory. JAMA pediatrics, 176(3), 224-225.

Ortony, A. , Clore, G. L. , & Collins, A. . (1988). The cognitive structure of emotions || reactions to agents. , 10.1017/CBO9780511571299(7), 134-155.

Qiu, Q., He, J., Chen, H., & Qiu, J. (2019, December). Research on the evolution law of emergency network public opinion. In 2019 12th International Symposium on Computational Intelligence and Design (ISCID) (Vol. 2, pp. 157-161). IEEE.

Ren, S., & Li, C. (2023). Public opinion communication mechanism of public health emergencies in Weibo: take the COVID-19 epidemic as an example. Frontiers in Public Health, 11, 1276083.

Rubdy, R. (2001). Creative destruction: Singapore's speak good English movement. World Englishes, 20(3), 341-355.

Scheu, A., Preuß, H., & Könecke, T. (2021). The legacy of the Olympic Games: A review. Journal of Global Sport Management, 6(3), 212-233.

Shaikh, M. A. M., Helmut, P., & Ishizuka, M. (2006, January). A cognitively based approach to affect sensing from text. In Proceedings of the 11th international conference on Intelligent user interfaces (pp. 303-305).

Shin, E. H. (2007). State, society, and economic development in sports life cycles: The case of boxing in Korea. East Asia, 24, 1-22.

Singhal, K., Agrawal, B., & Mittal, N. (2015). Modeling Indian general elections: sentiment analysis of political Twitter data. In Information Systems Design and Intelligent Applications: Proceedings of Second International Conference INDIA 2015, Volume 1 (pp. 469-477). Springer India.

Sun, X., Li, C., & Ren, F. (2016). Sentiment analysis for Chinese microblog based on deep neural networks with convolutional extension features. Neurocomputing, 210, 227-236.

Sznajd-Weron, K., & Sznajd, J. (2000). Opinion evolution in closed community. International Journal of Modern Physics C, 11(06), 1157-1165.

Thorpe, H., & Wheaton, B. (2019). The olympic games, agenda 2020 and action sports: the promise, politics and performance of organisational change. International Journal of Sport Policy and Politics, 11(3), 465-483.

Watts, D. J., & Dodds, P. S. (2007). Influentials, networks, and public opinion formation. Journal of consumer research, 34(4), 441-458.

Wan, B., Wu, P., Yeo, C. K., & Li, G. (2024). Emotion-cognitive reasoning integrated BERT for sentiment analysis of online public opinions on emergencies. Information Processing & Management, 61(2), 103609.

Wheaton, B., & Thorpe, H. (2018). Action sports, the Olympic Games, and the opportunities and challenges for gender equity: The cases of surfing and skateboarding. Journal of sport and social issues, 42(5), 315-342.

Wheaton, B., & Thorpe, H. (2019). Action sport media consumption trends across generations: Exploring the Olympic audience and the impact of action sports inclusion. Communication & Sport, 7(4), 415-445.

Wheaton, B., & Thorpe, H. (2021). Action sports and the Olympic Games: Past, present, future. Routledge.

Wheaton, B., Roy, G., & Olive, R. (2017). Exploring critical alternatives for youth development through lifestyle sport: Surfing and community development in Aotearoa/New Zealand. Sustainability, 9(12), 2298.

Xie, Q., Han, Q., & Chen, D. (2022). Analysis of Sports Popular Trend Based on Public Opinion Mining of New Media. Mathematical Problems in Engineering, 2022.

Yau, C. K., Porter, A., Newman, N., & Suominen, A. (2014). Clustering scientific documents with topic modeling. Scientometrics, 100, 767-786.

You, Y., Li, J., Reddi, S., Hseu, J., Kumar, S., Bhojanapalli, S., ... & Hsieh, C. J. (2019). Large batch optimization for deep learning: Training bert in 76 minutes. arXiv preprint arXiv:1904.00962.

Yu, B., & Ni, Y. (2024). Reveal the Characteristics and Commercial Value of Public Opinion Dissemination in the Big Data Era—Take the Internet Community Zhihu as an Example. In INTERNET FINANCE AND DIGITAL ECONOMY: Advances in Digital Economy and Data Analysis Technology The 2nd International Conference on Internet Finance and Digital Economy, Kuala Lumpur, Malaysia, 19–21 August 2022 (pp. 765-780).

Zhang, D., Xu, H., Su, Z., & Xu, Y. (2015). Chinese comments sentiment classification based on word2vec and SVMperf. Expert Systems with Applications, 42(4), 1857-1863.

Zhang, L., Wei, J., & Boncella, R. J. (2020). Emotional communication analysis of emergency microblog based on the evolution life cycle of public opinion. Information Discovery and Delivery, 48(3), 151-163.

Zhao, L., Alhoshan, W., Ferrari, A., Letsholo, K. J., Ajagbe, M. A., Chioasca, E. V., & Batista-Navarro, R. T. (2021). Natural language processing for requirements engineering: A systematic mapping study. ACM Computing Surveys (CSUR), 54(3), 1-41.

Zheng, J., & Liu, Y. (2023). What does Chinese BERT learn about syntactic knowledge?. PeerJ Computer Science, 9.

Zhou, Q., & Jing, M. (2020). Multidimensional mining of public opinion in emergency events. The Electronic Library, 38(3), 545-560.

Zhou, Q., Xu, Z., & Yen, N. Y. (2019). User sentiment analysis based on social network information and its application in consumer reconstruction intention. Computers in Human Behavior, 100, 177-183.

曾润喜,杜换霞 王君泽.(2014).网络舆情指标体系、方法与模型比较研究.情报杂志(04),96-101.

常曹育.(2019).基于机器学习的中文微博情感分类技术研究硕士学位论文,江苏科技大学).https://kns.cnki.net/kcms2/article/abstract?v=1fhXvtqifPKcZ8uybWUTzjXjFiQ1JlkCPik8bUXQhHI6Aq-Ra5tpmzB7sCDydORmVX711PsVU2kHnLhmPr18wOu30YPF-zyDkbEhnG2wkRAVNNMxTig6g0Qbu1ApEFgmxyqY_N6e1wcwUIQn3BWSWQ==uniplatform=NZKPTlanguage=CHS

陈虹 & 沈申奕.(2011).新媒体环境下突发事件中谣言的传播规律和应对策略.华东师范大学学报(哲学社会科学版)(03),83-91+154. doi:10.16382/j.cnki.1000-5579.2011.03.017."

陈晓美, 高铖, & 关心惠. (2015). 网络舆情观点提取的 lda 主题模型方法. 图书情报工作(21), 6.

陈子豪.(2020).碗式滑板世界格局及我国奥运备战对策研究.辽宁体育科技(04),115-118+142.doi:10.13940/j.cnki.lntykj.2020.04.025."

崔鹏,张巍,何毅 齐婧.(2018).突发公共事件网络舆情演化及政府应对能力研究.现代情报(02),75-83+95.

戴维·德·维利尔斯(Dawid J.de Villiers). (2009). 双赢的体育与旅游 win-win sports and tourism. 体育文化导刊(11), 153-154.

董光恒. (2008). 积极情绪与消极情绪启动对冲动控制影响的差异研究 博士 学位论文, 辽宁师范大学). https://kns.cnki.net/kcms2/article/abstract?v=S5uBaE2M3OevGTcrWc5pY3ss6bb255Vt1cgBLfw4TkxJSp8C-ckFHYPIL3lM7srOeZ4oN8OdoDvP0ZNt0DEr8pyqSUmiGYYiDOwUYYWgh51buaIAfUqzyN4ZIL50Thr3ophevF49u3FiZ6YpxV8Ykg==&uniplatform=NZKPT&language=CHS

杜洪涛,王君泽 李婕.(2017).基于多案例的突发事件网络舆情演化模式研究.情报学报(10),1038-1049.

杜李惠. (2014). 浅析初中化学教学中的情感教育. 课程教育研究:新教师教学, 000(025), 282-282.

范爽. (2019). 《人民日报》奥运赛事报道的话语变迁研究(1984-2016). (Doctoral dissertation, 河北大学).

冯江平,史俊洋,陈虹,赵舒贞 & 王鑫.(2012).网络社会事件发展过程中的舆情心理分析——以“躲猫猫”事件为例. 云南师范大学学报(哲学社会科学版)(04),60-70.

高莎莎. (2009). 《金瓶梅》在明清时期的传播与禁毁研究. (Doctoral dissertation, 中国海洋大学).

谷云峰.(2008).消极政治因素对北京奥运会的影响及应对策略研究.沈阳体育学院学报(02),28-30.

韩雨 张宇.(2022).信源、认知与情感：网络爱国主义生成机理及引导研究——基于 2010-2019 年网络舆情事件的模糊集分析.东南传播(01),32-36.doi:10.13556/j.cnki.dncb.cn35-1274/j.2022.01.001.

郝阳阳. (2019).《奥林匹克 2020 议程》三大支柱实施进展及问题研究. (eds.) 第十一届全国体育科学大会论文摘要汇编 (pp.662-663). 北京体育大学;
和立新.(2008).第十一届全运会与山东省经济社会互动的分析. 山东体育学院学报(06),14-16. doi:10.14104/j.cnki.1006-2076.2008.06.005.
何云.(2019).数据特性驱动的软件系统文本特征定位方法研究博士学位论文,云南大学).https://link.cnki.net/doi/10.27456/d.cnki.gyndu.2019.000044doi:10.27456/d.cnki.gyndu.2019.000044.
侯利君 张峰.(2018).微博突发事件负面舆情信息传播机理及应对策略.数学的实践与认识(23),167-174.
皇甫璐雯 & 毛文吉.(2017).一种基于 OCC 模型的文本情感挖掘方法.智能系统学报(05),645-652.
黄小波.(2016).夏季奥运会竞赛项目膨胀原因分析及对策研究.安徽体育科技(05),1-2+21.
黄一龙,李培峰 & 朱巧明.(2015).中文事件相关性语料库构建及识别方法.计算机工程与科学(12),2306-2311.
焦芳钱 宋志刚.(2023).现代夏季奥运会新增项目演变趋势下我国战略应对研究.(eds.)第十三届全国体育科学大会论文摘要集——专题报告(体育社会科学分会)(pp.336-338).厦门大学体育教学部;国家体育总局教练员学院;doi:10.26914/c.cnkihy.2023.066077.
金慧兰.(2014).网络舆情分析对体育管理工作的影响研究. 浙江体育科学(05),10-15.
康伟.(2012).突发事件舆情传播的社会网络结构测度与分析——基于“11·16 校车事故”的实证研究. 中国软科学(07),169-178.
李昌祖 张洪生.(2010).网络舆情的概念解析.现代传播(中国传媒大学学报)(09),139-140.doi:10.19997/j.cnki.xdcb.2010.09.033.
李闯.(2022).亚文化视角下中国青年滑板圈层文化的形成及发展机制.当代青年研究(06),48-58.
李峰, 王玲玲, & 杨菲. (2023). 一种基于深度学习的情感分析方法,系统及设备. CN115982355A.

李昊青,兰月新,侯晓娜 张琦.(2015).网络舆情管理的理论基础研究.现代情报(05),25-29+40.

李黎.(2017).基于文本挖掘的网络舆情情感倾向及演化分析硕士学位论文,湘潭大学).

李泽远, & 王健宗. (2021). 基于联邦学习的对抗生成网络模型训练方法及其相关设备. CN202110758657.4.

林崇德,杨治良,黄希庭.(2004).《心理学大辞典》前言.心理科学(05),1154.doi:10.16719/j.cnki.1671-6981.2004.05.033.

刘恒旺. (2018). 基于 OCC 模型和深度学习的网络舆情情感识别研究 硕士 学位论文, 南京理工大学). https://kns.cnki.net/kcms2/article/abstract?v=S5uBaE2M3OcWrXk6Mf_tbNCakU8L82l8jH4sI0wZWCkUZsV-iUxALMvxErBt9tcBhV-GnY9sgt-e3DSqIi0Bn0lg57AEo-acwPYsFXYtdoZQcUl00RGRmucxJ48E5Hird5NrnMbjAO_f8j_EzOiu1g==&uniplatform=NZKPT&language=CHS

刘继, & 顾凤云. (2022). 基于 bert 与 bilstm 混合方法的网络舆情非平衡文本情感分析. 情报杂志, 41(4), 7.

刘杰. (2012).网络民意及其对公共政策影响研究. (Doctoral dissertation, 天津师范大学).

刘茜 李倩.(2021-12-15).滑板攀岩冲浪或纳入 2028 洛杉矶奥运会小众运动如何破圈出道？.21 世纪经济报道,011.doi:10.28723/n.cnki.nsjbd.2021.005246.

刘世斌. (2019). 高校突发事件网络舆情演进规律及应对策略. 重庆广播电视大学学报, 31(5), 5.

罗梦珍.(2020).基于深度学习的医药专利自分类模型研究硕士学位论文,西华师范大学).

罗玉婷.(2023).网络不文明评论对受众情绪体验和传播行为的影响研究——基于认知评价理论视角.新媒体研究(15),43-46.

马钰.(2023).重大突发事件下微博舆情情感分析及演变趋势预测技术研究硕士学位论文,郑州轻工业大学).

庞娜.(2008).基于 OCC 模型的文本情感识别方法的研究硕士学位论文,太原理工大学).
https://kns.cnki.net/kcms2/article/abstract?v=sf24_f5fySaugO3fxR8njKToL2VNw4Z81_l3nFJuaageI7-n3Y_JwdYnTX2QikzCck8KHSXUUz0cTBJZBAaQGzNL324dmVSM3s3_NsHmEyRtrzuAdEm6iUfuc7f4lUXHymQpT2hY03UBFO8qMOOh6g==uniplatform=NZKPTlanguage=CHS

蒲毕文.(2014).基于社会网络分析的体育赛事舆情传播实证研究.山东体育学院学报(06),11-18.doi:10.14104/j.cnki.1006-2076.2014.06.025.

乔世君, & 张世英. (2000). 用 gibbs 抽样算法计算定数截尾时 weibull 分布的贝叶斯估计. 数理统计与管理, 19(2), 6.

邱洁. (2024)基于主题模型的社交媒体主题挖掘和文献影响力预测分析. (Doctoral dissertation, 西南大学).

沈昕怡 徐成龙.(2021).基于微博文本挖掘的体育赛事网络舆情研究——以东京奥运会为例.新闻研究导刊(23),149-151.

石峰. (2015). 青少年滑板运动爱好者运动损伤特征研究. 青少年体育(1), 2.

孙小龙 查建芳.(2017).网络体育信息传播的构成形态、舆论极化与引导机制研究.体育与科学(01),58-64.doi:10.13598/j.issn1004-4590.2017.01.008.

汤景泰. (2015). 新媒体语境下突发事件舆论引导中的整合议题管理. 中国应急管理(7), 5.

陶长春. (2014). 网络谣言对民意的表达与歪曲 博士 学位论文, 武汉大学).
https://kns.cnki.net/kcms2/article/abstract?v=S5uBaE2M3OchKL6sdYsHGhQ5ia6JwOYyqvX6JBNBXqhdKTccHS12EtX-oGqzythRwIJa1rGp4mZ4eVrIzxQ0QeZLeenU7QMQhna-svBVdlSZEtUPTOo-83_R-zpkofnV9--HygfSHqQDJS3gil8YEw==&uniplatform=NZKPT&language=CHS

王洪浩.(2019).中文分词算法在搜索引擎应用中的研究.中小企业管理与科技(下旬刊)(01),103-104.

王佳楠 梁永全.(2021).中文分词研究综述.软件导刊(04),247-252.

王来华.(2010).论舆情研究的两个需要. 天津社会科学(04),70-74.
doi:10.16240/j.cnki.1002-3976.2010.04.023."

王磊 邹薇.(2013).夏季奥运会项目设置的演变与价值取向.成都体育学院学报(06),33-37.doi:10.15942/j.jcsu.2013.06.002."

王美玲.(2023).我国小众体育兴起的要素驱动、现实困境及突破路径. 运动精品(10),27-29+33.

王宇超. (2022). 我国滑板运动项目市场化发展困境及对策研究. 当代体育科技, 12(4), 4.

王志强.(2023).数智时代体育突发事件舆论引导机制研究.(eds.)第十三届全国体育科学大会论文摘要集——专题报告(体育管理分会)(pp.733-735).江苏理工学院;doi:10.26914/c.cnkihy.2023.068913.

魏现辉.(2014).面向评论的跨领域情感分类研究硕士学位论文,大连理工大学).https://kns.cnki.net/kcms2/article/abstract?v=w1Je9LIFm5BFNsINdYqR3puMXkVSEfbKGUhDfwvaCBV0ouxxAek_vFofY5pebLT8rqHZBGJUIZS76qsiN6Z5JACZfuAxNA5rPjQF7riKrajiYFzn0CbEl20QdCBIpQiUdt7L4wWgyVghshrhGW_dDw==uniplatform=NZKPTlanguage=CHS

吴文峰,肖谋文,王鑫 & 刚伟.(2013).体育舆论及其引导策略刍论. 体育学刊(05),36-38. doi:10.16237/j.cnki.cn44-1404/g8.2013.05.012.

谢恬.(2023).基于微博话题的北京冬奥会体育赛事网络舆情研究硕士学位论文,成都体育学院).https://link.cnki.net/doi/10.26987/d.cnki.gcdtc.2023.000027doi:10.26987/d.cnki.gcdtc.2023.000027.

徐翔, 靳菁, & 吕伟欣. (2018). 网络舆情作为社会传感器对股票指数的影响———基于 lda 主题模型的挖掘分析. 财务与金融(6), 9.

杨畅,朱琳,兰月新 张慧玉.(2014).基于案例分析的群体性事件网络舆情传播规律及对策研究.现代情报(10),10-14+19.

杨沛君. (2018).网络亲社会行为中自主性、自我建构对接受者感戴情绪的影响. (Doctoral dissertation, 华中科技大学).

杨硕. (2014). 员工工作价值观和组织忠诚的代际差异及关系研究 硕士学位论文,河北大学).https://kns.cnki.net/kcms2/article/abstract?v=S5uBaE2M3OfNfrrOUe-bVzAPkjCRl1R1a_n1Pov5GAnS9VB2a2IAu91BWvsmL3Mw3v_bDwh0 jxCDPtEmFYdSQ5s_xLYLCAYL-HSUjAHXXLEBPsHvc17dG70WOfl4CLKHrQEt8xZ56RtcG1vzFVVrnA==&uniplatform=NZKPT&language=CHS

姚晓娜. (2009). BBS 热点话题挖掘与观点分析. (Doctoral dissertation, 大连海事大学).

叶正,傅灵,覃俊 & 刘晶.(2023).基于利用 BERT 不同层信息的微调策略的对话状态追踪.中南民族大学学报(自然科学版)(03),327-333.doi:10.20056/j.cnki.ZNMDZK.20230306.

由丽萍, 刘越, & 王世兴. (2022). 融合自注意力机制和语义词典的危机情绪分类研究. 情报理论与实践(005), 045.

袁泉.(2021).基于一维时序卷积神经网络的锂离子电池 SOH 估算硕士学位论文,长安大学).https://link.cnki.net/doi/10.26976/d.cnki.gchau.2021.001327doi:10.26976/d.cnki.gchau.2021.001327.

张丽娟 & 罗娟.(2012).网络舆情热点事件中的网民行为效应.青年记者(23),51-52.doi:10.15997/j.cnki.qnjz.2012.23.071.

张玉亮.(2012).突发事件网络舆情的生成原因与导控策略——基于网络舆情主体心理的分析视阈.情报杂志(04),54-57.

赵建兴. (2003). 生命周期理财理论与实践的新发展. 经济师(5), 2.

赵战胜.(2018).基于系统动力学的自然灾害信息传播研究硕士学位论文,中国地质大学(北京)).https://kns.cnki.net/kcms2/article/abstract?v=w1Je9LIFm5BJeivS7axvQPtOQHd3uZJVos8qVkD7APz7nZRkz2ncY4bzJh3XPOEqOtSD0JCMPSLEwanx5Io9-1AX2GqDzN9ZZPd27ySKHl3cRUG1Fbcg8lt4TnN_AeoXbXI5OCzXhrAqrzCA3WN8aQ==uniplatform=NZKPTlanguage=CHS

郑慧.(2011).浅议群体时代下网络舆论的非理性硕士学位论文,西南政法大学).

中国互联网络信息中心.第 51 次中国互联网络发展状况统计报告[R]. https://www.cnnic.cn/n4/2023/0303/c88-10757.html.

周昕,黄微,滕广青 李瑞.(2016).网络舆情传播模式解析与重构研究.情报理论与实践(12),25-30.doi:10.16353/j.cnki.1000-7490.2016.12.005.