Movie Audience Preference Mining and Content Generation Strategy Based on Big Data Analysis

Qianlin Liu

Liaoning Communication University, Shenyan, 110136, Liaoning, China

qianlin.liu@connect.polyu.hk

Keywords: Big Data Analysis; Movie Audience Preference; Content Generation Strategy; Theme Selection; Role-Building

Abstract: With the increasingly fierce competition in the film industry, it is very important to dig deep into audience preferences and formulate effective content generation strategies accordingly. This paper focuses on the research of movie audience preference mining and content generation strategy based on big data analysis. Through the comprehensive application of big data theory and the integration of multi-source data such as social media, online ticketing platform and video website, the preferences of movie audiences are deeply analyzed. With the help of big data analysis, we can have a clear insight into the audience's preferences in subject matter, role, plot, audio-visual effects and so on. For example, movies with different themes show dynamic changes in box office share and search popularity, and the audience has obvious preference for specific character characteristics. Based on these findings, this paper constructs a film content generation strategy theory covering theme and theme selection, role shaping and actor selection, plot structure and rhythm control, and visual and auditory effects. The research aims to help the film industry accurately grasp the needs of the audience, create more competitive works in the market and promote the high-quality development of the film industry.

1. Introduction

Driven by the wave of the digital age, the film industry is undergoing unprecedented changes. The rapid development of big data technology has provided a brand-new perspective and tools for the film industry, enabling it to gain an unprecedented insight into the preferences of movie audiences [1]. At present, the competition in the global film market is becoming increasingly fierce. How to accurately grasp the needs of the audience and create film works that meet the market expectations has become the key to the development of the film industry [2]. Traditional market research methods, such as questionnaires and interviews, are often interfered by sample size, geographical restrictions and subjective factors when obtaining audience preference information, and it is difficult to fully and accurately reflect the audience's true preferences [3]. With its massive data collection ability and efficient data analysis algorithm, big data analysis can capture the audience's behavior data on various platforms in real time and dynamically, including viewing history, ratings, comments, etc., which provides strong support for in-depth mining of audience preferences [4].

From the perspective of academic research, many scholars have discussed the preference of movie audiences, but most of them are limited to traditional research methods. With the wide application of big data technology, the research on movie audience preference mining and content generation strategy based on big data analysis has become a new academic hotspot [5]. The purpose of this study is to fill some gaps in this field, explore how to tap audience preferences with the help of big data analysis, and build a scientific and reasonable film content generation strategy based on this.

This study has important theoretical and practical significance. Theoretically, it enriches and expands the theoretical system of film industry and audience research, and provides new ideas and methods for further research by subsequent scholars. In practice, it is helpful for film producers,

  DOI: 10.25236/etmhs.2025.030

distributors and related practitioners to locate the market demand more accurately, optimize film creation, marketing and other links, enhance the market competitiveness of films and promote the sustainable development of the film industry. This study will innovatively integrate multi-source data, comprehensively use data mining, machine learning and other technologies, break through the limitations of single data source and simple analysis methods in the past, and strive to provide practical strategic suggestions for the development of the film industry.

2. The theoretical basis of movie audience preference mining

Movie audience preference mining relies on the theory of big data and audience preference. Big data, as a massive, high-growth and diversified information asset, has a wide range of sources, covering social media, online ticketing platforms, video websites and so on [6]. These platforms record the audience's viewing behavior, evaluation, discussion and other data, which constitute the basic material for mining audience preferences. The 4V characteristics of big data, namely Volume, Variety, Velocity and Value, provide the possibility for in-depth analysis of audience preferences [7]. Massive data ensures that the behaviors of various audiences can be captured; Various data types, such as text, images, behavior data, etc., reflect the audience's preferences from multiple dimensions; High-speed data update enables the analysis to keep up with the changes of audience preferences; The value of data can be reflected through effective mining and analysis.

The formation of audience preference is influenced by many factors, involving psychology and sociology. From the psychological point of view, the audience's personal interests, emotional needs, cognitive styles and so on affect their choice of movies [8]. For example, an adventurous and exciting audience may prefer action and science fiction movies; The audience with delicate emotions and the pursuit of resonance may have a special liking for plots and literary films. In terms of sociological factors, social and cultural background and group influence play a role. Under different cultural backgrounds, the audience's acceptance of film themes and themes is different; At the same time, individual movie-watching choices will also be influenced by families, friends and other groups [9]. These theories provide a framework for understanding audience preferences. Combining big data theory with audience preference theory, we can use big data technology to extract information related to audience preference from massive and complex data.

3. Theoretical construction of movie content generation strategy based on big data analysis

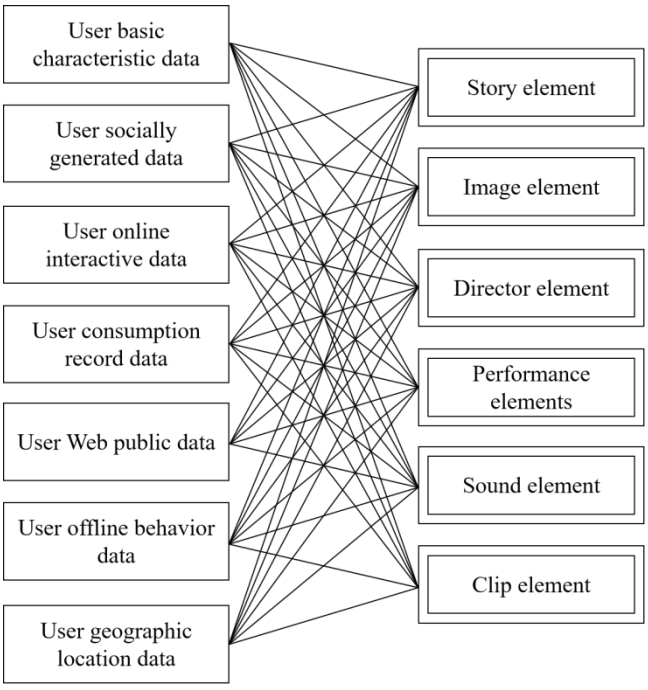

Figure 1 Correlation between user data and movie elements

In the era of big data, in order to achieve sustainable development, the film industry must accurately grasp the needs of the audience and formulate scientific and reasonable content generation strategies [10]. Audience preferences mined based on big data analysis can provide multi-dimensional guidance for movie content generation. The correlation between user data and movie elements is shown in Figure 1.

(1) Theme and theme selection strategy

Through the analysis of multi-source big data such as the popularity of social media discussion, search records of viewing platforms and box office data, we can clearly understand the popularity of different themes and themes among the audience. Table 1 shows the changing trend of popular themes such as science fiction, comedy and action in box office share and internet search popularity. As can be seen from the data in the table, the proportion of box office of sci-fi movies has been on the rise in the past five years, and its search popularity has also continued to rise. This shows that the audience is paying more and more attention to the future world, cutting-edge technology and other elements displayed by science fiction themes.

Table 1: Box Office Share and Search Popularity of Films of Different Genres in the Past Five Years

Year	Sci-Fi Genre	Search Popularity Index	Comedy Genre	Search Popularity Index	Action Genre	Search Popularity Index	Romance Genre	Search Popularity Index	Mystery Genre	Search Popularity Index
	Box Office Share		Box Office Share		Box Office Share		Box Office Share		Box Office Share	
2018	14.8%	122	21.9%	128	20.2%	123	17.7%	108	9.6%	98
2019	18.3%	134	19.7%	123	18.8%	120	16.6%	103	10.5%	102
2020	21.7%	152	17.6%	118	16.9%	116	15.4%	98	12.4%	107
2021	24.6%	163	15.8%	113	14.7%	112	13.9%	93	14.5%	112
2022	27.9%	178	14.1%	108	12.9%	107	11.8%	88	17.5%	117

Based on the results of such big data analysis, filmmakers should give priority to popular themes, such as science fiction and suspense, and pay attention to the integration and innovation of themes. For example, science fiction and comedy elements are combined to create imaginative and humorous stories to attract a wider audience. At the same time, we should dig deep into the theme connotation, and combine the current social hotspots, cultural trends and the audience's deep emotional needs, such as the concern for environmental protection and human exploration, so that the film is not only entertaining, but also can arouse the audience's thinking and emotional resonance.

(2) Role-building and actor selection strategy

The role is one of the key factors to attract the audience. Big data analysis can reveal the audience's preference for roles from the audience's evaluation of different role types and the popularity of role-related topics. Taking Table 2 as an example, the data shows that the characters with brave, intelligent and sense of justice are loved by the vast majority of the audience.

Table 2: Survey on Audience's Preference for Different Character Traits

Character personality traits	Preference Level (%)
Brave	84.5
Wisdom	79.3
Rich in a sense of justice	81.6
Humorous and witty	74.8
Fortitude	77.7
Gentle and kind	69.5

In terms of role-building, film creators should create a role image with distinctive personality and in line with the audience's psychological expectations according to these preferences. Creators should pay attention to the growth and transformation process of character images, and make them more three-dimensional and realistic through detailed portrayal. In terms of actor selection, big data can provide information such as the popularity index of actors, the characteristics of fan groups, the performance of past works and the audience's evaluation of their acting skills. Giving priority to actors who have a high degree of fit with the role image, high popularity and good reputation can attract the actors' own fans to watch movies, and at the same time, enhance the charm and

credibility of the role and enhance the attraction of the film with the help of the actors' performance ability.

(3) Plot structure and rhythm control strategy

The plot structure and rhythm of the film directly affect the audience's viewing experience. Through the big data analysis of the real-time feedback data of the audience in the process of watching movies, such as the pause and playback records of the video platform, and the plot-related evaluation in the film review, we can understand the audience's attention and acceptance of the plot development. Generally speaking, the audience tends to plot with compact plot, reasonable conflict and logic. In the plot structure, we should avoid procrastination and lengthy preparation, quickly introduce core conflicts, and constantly set suspense and reversal in the development of the story to maintain the audience's curiosity and attention. In the aspect of rhythm control, the rhythm is adjusted reasonably according to the tension and emotional atmosphere of the plot. Directors should accelerate the narrative rhythm of action scenes and key plot points to create a tense and stimulating viewing atmosphere; In the delicate place of emotional expression and role portrayal, the rhythm should be slowed down appropriately, so that the audience can have time to immerse themselves and feel the emotional changes of the role.

(4) Visual and auditory effect strategies

With the improvement of the audience's viewing level, the requirements for the visual and auditory effects of movies are getting higher and higher. Big data analysis can understand the audience's aesthetic preferences through the audience's evaluation of different movie screen styles, special effects quality and soundtrack sound effects. The film production team should choose appropriate visual styles based on the theme and story background of the film in terms of visual effect design. For example, fantasy themes are suitable for using colorful visual expression techniques, while artistic themes are more suitable for simple and elegant visual presentation. At the same time, it is necessary to ensure the quality and realism of the special effects production, and avoid affecting the audience's viewing experience due to rough and distorted special effects. In terms of auditory effects, it is necessary to carefully select music works that complement the development of the plot, using music to create an atmosphere and enhance emotional expression. The design of sound effect is also very important. It is necessary to ensure the clarity and three-dimensional sense of the sound, so that the audience seems to be there.

To sum up, based on big data analysis, the film content generation strategy is constructed, covering many key aspects such as theme, role actors, plot rhythm and audio-visual effects. Through in-depth excavation and rational use of big data, the film industry can more accurately meet the needs of the audience, create more attractive and market-competitive film works, and promote the high-quality development of the film industry.

4. Conclusions

This paper focuses on the movie audience preference mining and content generation strategy based on big data analysis, and has achieved a series of theoretical and practical results.

On the theoretical level, the study clarified the key role of big data in the mining of movie audience preferences. With its mass, diversity, high speed and value, big data provides a rich and real-time data foundation for in-depth understanding of the audience. Combining the theory of audience preference formation in psychology and sociology, through big data analysis, we can analyze audience preferences from multiple dimensions, such as tracking the popularity of movies with different themes and gaining insight into the preferences of characters, which enriches the theoretical system of film industry and audience research. In terms of practical strategy construction, a series of film content generation strategies based on the results of big data analysis provide a practical direction for the development of the film industry. In terms of theme and theme selection, according to the theme heat and social and cultural trends reflected by big data, it will help producers accurately grasp the market demand and realize the innovation and integration of themes. Role shaping and actor selection strategies, with the help of big data to consider role preference and actor fitness, can enhance the role appeal and movie appeal. The plot structure, rhythm control

strategy and visual and auditory effect strategy, starting from improving the audience's viewing experience and guided by big data analysis, make the film more suitable for the audience's expectations in content presentation and sensory experience.

Although big data provides abundant information, factors such as data quality, limitations of analysis methods and dynamic changes of audience preferences may affect the accuracy and timeliness of research results. Future research can further optimize the data collection and analysis methods, strengthen the continuous tracking of the dynamic evolution of audience preferences, and continuously improve the movie content generation strategy based on big data analysis. Overall, this study provides a useful reference for the development of the film industry in the era of big data, and is expected to promote the film industry to better meet the needs of the audience and achieve sustainable development.

Acknowledgements

This paper is a research outcome of the 2024 Basic Scientific Research Project of the Liaoning Provincial Department of Education, titled "Research on the Mainstream Development Strategies of Micro-short Dramas under the Concept of Integration of Sports, Culture, and Tourism" (Project Number: LJ112413957011).

References

[1] Zhang Jianzhen, Wu Haiqing. New Technology Ethics: Cinema and the Fourth Industrial Revolution[J]. Film Art, 2024(5):53-60.

[2] Zhang Huifeng, Jing Tian. Research on Evaluation Methods of Film Consumption Demand Based on Consumers' Rational Decision-Making[J]. Journalism Bimonthly, 2023(8):101-115.

[3] Lan Duo, Zhu Yicheng. A Framework Study on the Overseas Online Communication of Chinese Films in Cross-Cultural Context[J]. Contemporary Cinema, 2022(8):75-81.

[4] Zhao Shiyue, Zhou Tao, Han Xiaopu, et al. Analysis of Gender Differences in the Market Influence of Chinese Film Actors[J]. Journal of University of Electronic Science and Technology of China, 2024, 53(2):271-276.

[5] Li Guangming, Yang Panpan, Gu Chan. A Film Recommendation System Based on Flink for Dynamically Perceiving User Interest Drift[J]. Chinese Journal of Electron Devices, 2024, 47(5):1425-1433.

[6] Zheng Gencheng. Application of Artificial Intelligence Technology in Film and Television Creation and Its Ethical Examination[J]. Journal of Beijing Film Academy, 2024(11):75-81.

[7] Wang Gang, Xiao Ke. Influence of Internet Culture on Film Marketing in the New Media Context and Strategy Optimization[J]. Movie Literature, 2025(3):135-139.

[8] Xu Wei. Innovation Paths for In-depth Integration of TV Media Empowered by Big Data[J]. Media, 2023(19):33-35.

[9] Wu Ting. Mining and Analysis of Typical Opinions in Animated Film Reviews in the Context of Big Data[J]. Movie Review, 2022(20):11-16.

[10] Sun Changzhi, Yi Mingyang, Jiang Jingjing, et al. Research on Sentiment Analysis of Film Short Review Data Based on Deep Topic Autoencoder Model[J]. Information Science, 2024, 42(11):69-75.