

AUDIENCE SATISFACTION PREDICTION OF CHINESE WEB DRAMA BASED ON BIG DATA TECHNOLOGY AND SUPPORT VECTOR MACHINE

Pei Cao^{1*}, Jamilah Binti Jamal²

¹Ph.D candidate, School of Multimedia Technology and Communication, College of Arts and Sciences, Universiti Utara Malaysia, UUM Sintok, Kedah Darul Aman, Malaysia, 06010

²Dr., Senior Lecturer, School of Multimedia Technology and Communication, College of Arts and Sciences, Universiti Utara Malaysia, UUM Sintok, Kedah Darul Aman, Malaysia, 06010

Received: 22 January 2024

Accepted: 25 May 2024

First Online: 30 June 2024

Research Paper

Abstract: *With the explosive growth of online video content, web dramas have become an important component of entertainment consumption for a large audience. This article constructs a prediction model that combines big data processing technology and Support Vector Machine (SVM) algorithm. This model utilizes rich historical data features such as drama genre, cast, production costs, and promotional work to predict future audience satisfaction with online dramas. By optimizing model parameters and feature selection, analysis of monthly audience satisfaction data over the past year revealed that different types of dramas have a significant impact on audience satisfaction, with comedy genres consistently scoring above 85 points on average per month. The average monthly satisfaction score for ancient costume dramas is about 82 points. In contrast, the average monthly satisfaction score for thriller and suspense dramas is about 78 points. This indicates that there are different evaluations of drama due to personal preferences. War-themed TV dramas, due to their more serious themes, have a relatively stable audience, with an average monthly satisfaction score of about 75 points. This study constructs a predictive model that can effectively improve prediction accuracy and provide scientific decision support for film and television producers, investors, and platform operators. Conversely, this can help them better grasp market trends and create high-quality works that better meet audience expectations.*

Keywords: *Big Data Technology; Support Vector Machines; Audience Satisfaction Predictive model; Web Drama; Sentiment Analysis*

*Corresponding Author: cao_pei@ahsgs.uum.edu.my, (P. Cao),
jamilah@uum.edu.my, (J. B. Jamal)

1. Introduction

China Internet Network Information Centre On February 25, 2022, the 49th Statistical Report on Internet Development in China was issued by the China Internet Network Information Centre, pointing out many remarkable changes that have taken place in China's digital development. By the end of 2021, the number of Chinese internet users had reached 1.032 billion, while the internet penetration rate was 73.0% (Alsayat & Ahmadi, 2023). In China, nearly all internet users-99.7%-access online content through a mobile device, indicating the core role of online video in everyday life and entertainment. As such, with the increasing demand for high-quality content, the video platforms are compelled to increase their standards of production, thus moving forward with the development of the film and television industry in China. Works like "The Hidden Corner" and "The Silent Truth" set new standards for high-quality serials. The reintroduction of iQiyi's "Mist Theater" with "The Mystery of the Octagonal Pavilion" this October underlined the ongoing popularity of suspense dramas (Chen et al., 2022).

It is an undeniable fact that domestic online suspense dramas have set off a climax in production, broadcasting and ratings (Chen et al., 2022). In recent years, the number of web suspense dramas has obviously increased. During the process of the audience's preference for genres becoming more specialized and diversified, including several elements, the suspense drama genre became an important part of the web drama market. Through the investigation of the audience's motives in viewing online suspense dramas and the satisfaction of the platforms, such as "Mist Theater," this present research will analyse the composition of audiences to objectively understand the reality of the genre and provide support for future online suspense drama development. At the same time, an examination regarding the relationship between the demand and satisfaction of audiences can enrich the empirical application of Pamgrin's theories related to expected satisfaction (GS) and actual satisfaction (GO). Therefore, the understanding of the needs and satisfaction of the audience will be quite informative for further development and distribution of online suspense dramas (Dudzik et al., 2024; Simonov et al., 2023).

It has become an urgent problem in industry: how to accurately estimate the satisfaction of the audience and optimize the production to make products more competitive. This paper intends to use big data technology and the support vector machine method to develop a predictive model for audience satisfaction. The paper is based on detailed analysis of viewing data and assessment information of web drama viewers. It will provide scientific and effective decision support to producers in the web drama industry (Shmueli & Polak, 2024). This research would not only benefit the development and innovation within the web drama industry but also could provide informative insights and references for audience research across other cultural sectors.

2. Related Works

2.1 Research Significance

This paper uses big data technology and support vector machine, takes Mist Theater as the research object, explores the reasons for the popularity of online

Audience Satisfaction Prediction of Chinese Web Drama Based on Big Data Technology and Support Vector Machine

suspense dramas, the audience's viewing behaviour, and the audience's demand satisfaction, which will enrich the theory of "use and satisfaction" in the field of audiences. Research is also beneficial to promoting the perfection and development of the dissemination of online suspense dramas. Currently, Internet suspense dramas are a popular form of mass entertainment and thus deserve equal academic attention. Scholars both at home and abroad have researched the aesthetic and cultural phenomenon of "Internet suspense dramas," as well as their methods of narration, plot structures, and operation models. From the review of current literature, it can be seen that there is still more to research, particularly in terms of the motivation for audience demand and satisfaction (Aziz et al., 2022; Jingmiao, 2024; Khetarpal et al., 2024). Therefore, empirical analysis and demand theory are combined in this paper to analyse, from the audience's point of view, the expected satisfaction when watching online suspense dramas and the satisfaction obtained after watching the fog theatre, and observe the impact of the satisfaction expected and obtained on the viewing behaviour of online suspense dramas. The following research results will be references for the progress of online suspense dramas in theory.

The escalating popularity of online suspense dramas has transformed numerous productions into phenomena, underscoring the importance of understanding and catering to this burgeoning audience segment (Li et al., 2023). It can determine the main groups of online suspense drama audiences and give aid to create some film and television companies. Thirdly, this research has significance for the audience on how they can meet their needs by contacting media resources. With the rapid development of the Internet technology, more abundant and varied opportunities can be used by the audience. In such a background, the awareness of self in audiences is clear, and it is exceedingly important for the audience to make choices of media content on their own needs. The development trend of online suspense dramas is bound to be based on the changing aesthetic needs of the audience and the general environment of policy regulation, and online suspense dramas should also assume social responsibilities, echo the emotions of the times, and respond to the emotional needs of the audience (Gordillo-Rodriguez et al., 2023). Its future content creation needs to be cultural sexuality and values, close to realistic themes. This article aims to improve network suspense dramas' quality and make them more align with audience preference. Based on empirical research insight and data, it is expected that the audience will be able to enjoy better quality and satisfactory feelings. At the same time, major video platforms can use the information provided to choose and spread superior popular culture that conforms to audience preference so as to create more economic benefits. The findings and conclusions drawn from this study will go a long way in ensuring that dramatists and distributors are guided to continue growing and developing network suspense dramas within the country.

2.2 Research on Web Drama

Generally, compared with other types of dramas, suspense dramas face the most serious and strict forms of censorship in many countries. Big data comes into play with the monitoring of real-time audience sentiment through social media. Moreover, the rise of the YouTube platform as a primary channel for content delivery has sparked a quest to understand underpinning factors of viewership and audience engagement tied to social media dynamics and sentiment analysis (Nieto-Ferrando et al., 2023). In a study of 16,278 images from 137 brand channels, Jang et al. (2023) found that

thumbnail design, brand reputation, and channel size are highly influential factors on video views. These findings dovetail with general trends in audience sentiment analysis wherein visual and reputational cues have influenced online engagements. Moreover, their study adds to this body of knowledge about ways in which brands may use social media and sentiment-driven strategies to drive performances from their contents. Along with the emergence of web dramas, the regulatory framework that controls online content is considerably slack; therefore, the development of online suspense dramas has seen a sudden increase. A review of the existing literature shows that domestic researches on network suspense dramas mainly focus on the causes of its popularity, strategies of narration, audience analysis, and the effects of communication (Liang & Zhang, 2022). Figure 1 shows the overview diagram of the research process of audience satisfaction prediction of Chinese web dramas, which shows the process and key links of the whole research, including the steps of problem definition, method selection, model construction, experimental verification and result analysis.

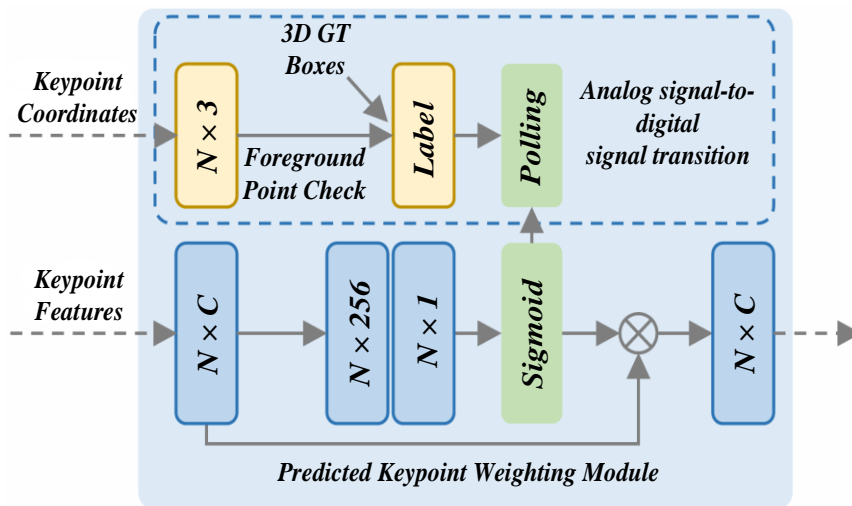


Figure 1: Overview Chart of the Research Process of Audience Satisfaction Prediction of Chinese Web Dramas

The Figure 1 is intended to help researchers systematically understand the whole process of satisfaction prediction and provide guidance for subsequent studies. From the perspective of relevant research on the cause of formation as the entry point, online suspense dramas, as a once niche genre drama, have achieved circle-breaking communication in the rapid development of recent years, setting off a climax of production, broadcasting and ratings, and have both popularity and word-of-mouth, the market performance is outstanding, and scholars mostly discuss the reasons for its successful exit from the circle. The Table 1 indicates that the web drama viewership records, comment information, social media responses, etc. will be integrated into one statistical database. It also contains comprehensive and extensive data which will be useful in the following analysis to forecast audience satisfaction. From the perspective of narrative as the entry point of relevant research, this aspect of the data mostly focuses on the analysis of the film and television text in the network suspense drama.

Audience Satisfaction Prediction of Chinese Web Drama Based on Big Data Technology and Support Vector Machine

The narrative strategy of suspense short plays is analysed from the perspective of detailed display (Mathur et al., 2023).

Table 1: Statistical Data Collection

Data Type	Volume of Data (Bar)	Data Source
Viewing Record	10,000,000+	Video Platform API
User Reviews	5,000,000+	Comment System
Social Media Interaction	20,000,000+	Social Media Platforms

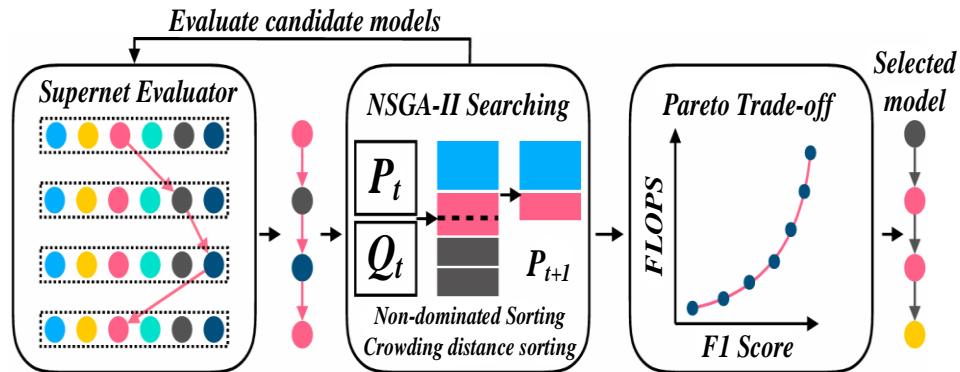


Figure 2: Flow Chart of Audience Data Analysis of Web Drama Based on Big Data Technology

Figure 2 shows the flow chart of audience data analysis of network dramas based on big data technology. This chart shows the complete process from data collection, storage, processing to analysis, highlighting the key role of big data in the deep mining of audience behavior mode. From the perspective of the audience as the entry point of relevant research, the continuous emergence of excellent online suspense dramas is inseparable from the audience who grew up in an open and diverse environment of the Internet, but there are few such research results (Moshawrab et al., 2023). Table 2 shows the results of feature extraction and pre-processing. Further, the most valuable features extracted from raw data are cleaned and standardized to present quality input data for further training of the models. Taking the perspective of the effect of communication as an entry point, most content on the internet in suspense dramas involves factors like murder, bloodshed, violence, etc., and these factors possess their own traffic and social topicality. Online suspense dramas are gradually being turned from perceptual satisfaction into realistic guidance. Improve the communication power of Internet suspense dramas by including American Hollywood film and television culture, combining it with traditional excellent Chinese culture, and adding national characteristic content to the story framework under the text transplantation.

Table 2: Feature Extraction and Pre-Processing Results

Feature Category	Number of Features	Example Features
Plot-Related	10	Plot Complexity, Plot Coherence, etc.
Actor Related	8	Actor's Acting Skills, Actor's Popularity, etc.
Making Related	7	Picture Quality, Sound Quality, etc.
Word-of-Mouth Communication	5	Number of Comments, Shares, etc.

Furthermore, recent advancements in mobile network communications and e-commerce have led to the birth of web series. It is thus sensible to scrutinize the video view prediction model. An advanced stacking ensemble model was proposed by (Mi et al., 2022), which introduced the sentiment scores coming from comments of viewers, which were computed using the SO-PMI algorithm. The results showed that including sentiment scores gives higher prediction accuracy compared to models without using sentiment scores, by using four algorithms: Random Forest, Gradient Boosting Decision Tree, Extreme Gradient Boosting, and Light Gradient Boosting. This makes evident the benefit that sentiment analysis is offering to enrich the view count predictions for web content. In addition to this, maintaining large digital platforms has led to big data being one of the major tools that are used in determining audience behaviour. For example, (Gutta et al., 2024) focused on the concept of big data analytics and machine learning models developed together to improve the accuracy of audience behaviour predictions. Moreover, these kind of approaches interrelated with web drama sites which create enormous web viewer data in the form of watch times, genre preferences, social media engagement, and others that can be mined in the sake of predicting satisfaction and engagement trends.

3. Methodology

3.1 Concept Definition and Dissemination Process of Domestic Web Dramas

The realm of film and television works is a constantly evolving and innovative landscape, where genres continually redefine themselves. This chapter meticulously defines the core concepts, delves into the dissemination process of domestic online suspense dramas, and provides insight into the evolution of these dramas within the broader context of suspense genre dissemination (Shardaghly, 2024; Zhao et al., 2024). Firstly, the explanation of the concepts and implications involved, so that it sets a foundation for a rigorous and thorough analysis. This section is dedicated to explaining the three main concepts: suspense drama, domestic online suspense drama, and "Mist Theater."

Table 3 presents an analysis of the factors influencing audience satisfaction. Although several researches have been conducted on the definition of suspense drama in China, a consensus has yet to be reached. Suspense drama is a kind of drama that uses suspense as the propellant for contexts and possesses advantages that no other type of drama has in narration. In this paper, suspense is defined based on the analysis of the concept "suspense" and the related research. Not found in "Ci Hai" or sources of word origin after searching, but one term most closely related is "suspense," which "Ci Hai" defines as the psychological tension while experiencing a literary or artistic work. Directors use all kinds of methods to build up suspense and raise the curiosity of the audience (Akbar et al., 2024). It meticulously crafts an overall narrative that teases the audience with lingering doubts, weaving suspense throughout the entire narrative arc. Through intricate plot twists, bizarre storylines, meticulous and intricate logical reasoning, as well as complex and multifaceted narrative techniques, it meticulously unfolds the entire detective process. This immersive experience keeps the audience engrossed alongside the protagonist, shrouding them in a veil of mystery akin to a dense fog, where the proverbial "Mount Tai" - a clear sense of direction or resolution - remains tantalizingly out of sight.

Audience Satisfaction Prediction of Chinese Web Drama Based on Big Data Technology and Support Vector Machine

Table 3: Analysis of Influencing Factors of Audience Satisfaction

Influencing Factors	Degree of Impact	Suggestions for Improvement
Drama Quality	High	Strengthen Script Creation and Enhance Plot Appeal.
Actor Performance	Medium	Selecting Excellent Actors and Improving Acting Skills.
Fabrication Quality	Medium	Invest More Resources to Improve the Production Level of Pictures and Sound Effects.
Word-of-Mouth Communication	Low	Strengthen Social Media Marketing and Expand Word-of-Mouth Influence.

Table 4 shows analysis of factors influencing audience satisfaction of web dramas. The table 4 systematically sorts out the key factors affecting the satisfaction of the audience, and provides an important reference basis for the producers to deeply understand and improve the works. It can be seen that the classification criteria of suspense dramas are mainly their narrative characteristics with a lot of suspense. This paper defines suspense drama as a type of TV drama or web drama in which detective protagonists solve puzzles and use reasoning methods to uncover cases or numerous mysteries in intricate plots.

Table 4: Analysis of Factors Influencing Audience Satisfaction of Web Dramas

Influencing factors	Describe	Degree of Impact (1-5)
Plot Attraction	The Innovation, Logic, Emotional Resonance of the Plot, etc.	4.5
Actor Performance	The Starring Acting Skills, Image and Character Fit.	4.0
Fabrication Quality	Production Level of Picture Effects, Sound Effects, Editing, etc.	3.8

3.2 Domestic Web Dramas

This paper mainly starts with the two key words of web drama and suspense drama, and uses big data technology and support vector machine technology to try to frame the concept of network suspense drama. Figure 3 shows the flow chart of the application of the support vector machine in the audience satisfaction prediction, mainly showing how to use the SVM for the satisfaction prediction process. Centred on the core steps of SVM algorithm training, this process involves parameter adjustment and prediction application. Leveraging the robust classification capabilities of SVM, it achieves precise predictions of audience satisfaction, offering significant support for the production and operation of web dramas. The broadcasting channel of web dramas is the Internet and produced by Internet companies and streaming media platforms, while the broadcasting channel of TV dramas is TV, which is the most essential difference between the two. Based on the analysis from the literature, the definition of domestic online suspense dramas in this paper is: the producer is a video platform or it cooperates with film and television companies, the narrative technique sets up more suspense, the themes include crime, science fiction, tomb robbery, etc, and solves the puzzles through scientific reasoning to find out the truth, so that the audience will have emotional reactions such as psychological tension and stimulation during the viewing process. The drama series is mainly broadcast on the Internet (Ponmalar & Dhanakoti, 2022).

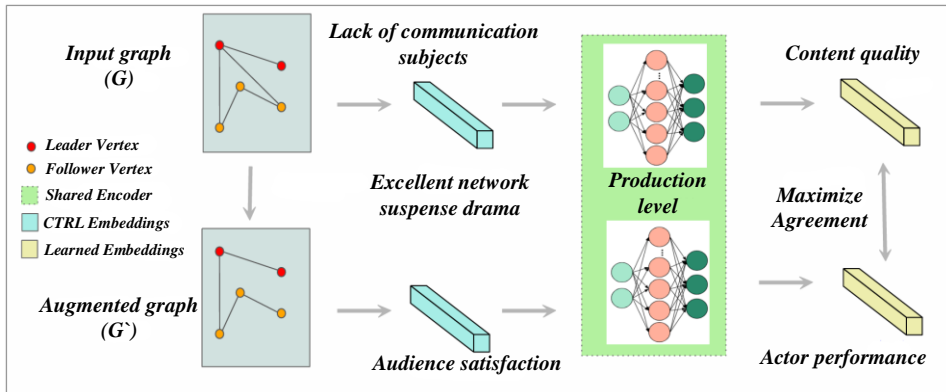


Figure 3: Flowchart of the Application of Support Vector Machine in Audience Satisfaction Prediction

3.3 Audience Communication Security and Privacy Protection

In carrying out audience satisfaction research on web dramas, one has to ensure that no infringement on privacy rights occurs in the data collection and analysis processes, but also has to safeguard the communication against unauthorized access and leakage. It's not only about ethics or legitimacy in the research conduct but also about audiences trusting the process and results of the research. Therefore, during the research process, a series of measures have been adopted to ensure the communication security and privacy protection of the audience. First, this article strictly abides by users' consent and minimisation of data principles. Data collection for audience does not happen unless and until comprehensive privacy policies have been shown to the audience concerning the purpose of data collection, the scope, and intent for which data are going to be used or shared. Consent from the audience will be explicitly requested before any collection of data commences, and obtained data is limited in scope to minimum; thus, only the information that is needed for the research shall be collected. On the other hand, for those personal sensitive information that needs to be collected, we have adopted separate encryption and storage while setting strict permission control on the access. In this way, it can be guaranteed that no unauthorized third party can access or tamper with such information.

Secondly, this article adopts advanced encryption technology and security protocols to protect the security of network communication and data transmission. This work utilize encryption protocols such as HTTPS to ensure the confidentiality and integrity of network communication, preventing data from being intercepted or tampered with during transmission. At the same time, special processing has been carried out for the transmission of sensitive information, prioritizing secure communication channels such as virtual private networks (VPNs) to further improve the security of data transmission. The process construction of data security and privacy protection has been strengthened. This article establishes internal data processing standards and operational procedures that comply with industry and regulatory requirements, clarifies the responsible persons and supervisory agencies for data processing, and ensures the legal, legitimate, and necessary use of data. At the same time, we have implemented strict permission control policies to restrict access and operation permissions to data, preventing unauthorized access and data leakage.

Audience Satisfaction Prediction of Chinese Web Drama Based on Big Data Technology and Support Vector Machine

Practices of transparency were adopted, and several auditing tools were used to locate and fix various security vulnerabilities as soon as possible. These various tools help us closely watch and control the flow of data and communication in our network for data integrity and security. This work follows an aggressive schedule for the backup and recovery testing of our data, which serves as a strong defence mechanism against sudden data loss or corruption. This leaves us with the very important concern of employee training because it is an important part of commitment to engage team members with increased awareness of data security practices. Regular training and educational activities have significantly enhanced awareness and the importance of data security and privacy protection among employees, who are now better able to observe relevant regulations and operating procedures that guarantee effective audience data security and privacy protection.

3.4 Connotation and Characteristics of Big Data

According to Laney's definition of big data, these digital objects and data that are collected and created through non-traditional legal deposits exhibit three key "3V" characteristics: volume, velocity, and variety. Volume refers to the massive scale of big data, where the level of data processed reaches the tera- or petabyte level. Velocity signifies the ever-increasing speed of big data technology in handling, processing, and analysing data. Variety represents the diverse range of data types that are processed in big data operations. Variability and effectiveness refer to the ability to ensure the quality and reliability of data, and value represents the ability to effectively transform big data into value or knowledge that can be beneficial for development (Razzak et al., 2020). The technical architecture and technical process of mining can be observed in (Basyouni et al., 2024). Regarding the connotation and characteristics of big data, this paper describes it as four points: information, technology, method and impact.

Table 5 shows SVM model parameter tuning result table. Moreover, Table 5 gives a different performance for different parameter combinations and provides data support to find the best parameters to ensure that the model performs best in predicting audience satisfaction. The first big-data acceleration is the continuous creation of data and enhancing data sharing and use (Selmy et al., 2023). Digitization, the transition from analogue signals to digital signals, prevailed in the 1990s. After the advent of the first commercial optical symbol recognition tool, large-scale digitization has also boarded the train of rapid development. Large-scale digitization is to transform traditional databases into machine-readable files (Alemdar & Codur, 2024). For example, Google in 2004 digitized over 15 million books collected from various universities such as Harvard, Stanford, and Oxford. Once a signal has been converted into a digital format, it can then be organized into a more structured dataset, often referred to as 'digitization.' With this, a researcher would be able to take advantage of a dataset retrieved from books to determine knowledge ranging in fields such as linguistics to sociology to historical epidemiology. Some view information as a manifestation of data, intangible assets that can create value for an enterprise; it is, therefore, inferred that information is the basic feed for big data.

Table 5: SVM Model Parameter Tuning Result

Parameter	Before Tuning	After Tuning
Kernel Function	RBF	Linear
Penalty Coefficient C	1.0	0.5
Loss Function ϵ	0.1	0.05

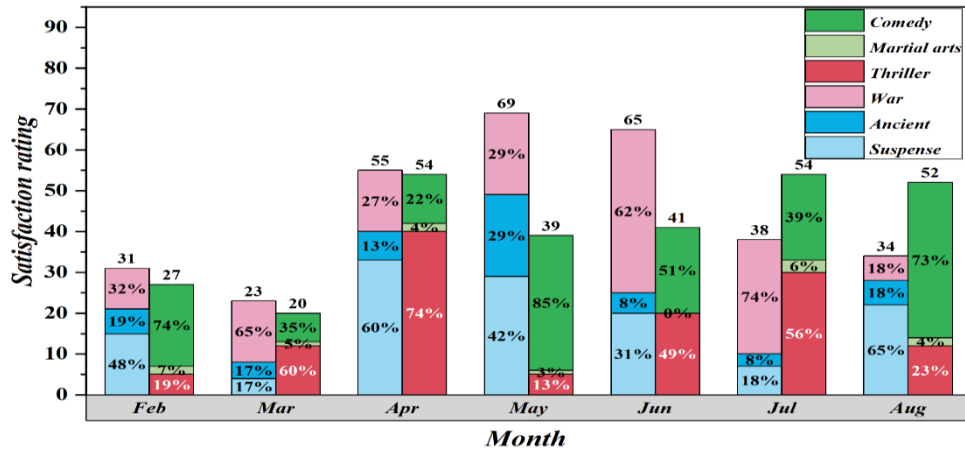


Figure 4: Plot Type and Satisfaction Relationship Diagram

Figure 4 shows the type of plot and the satisfaction of the audience. As the type of plot has a great impact on audience satisfaction, one may note in the given monthly data that comedy managed to keep the satisfaction scores above 85 points throughout the past year. Costume dramas also have a relatively high satisfaction average of about 82 points, supplemented in some months by the release of some high-quality new productions. In addition, thriller and suspense dramas are more volatile in satisfaction, averaging about 78 points, as the audience is attracted by specific plots. The serious theme of these dramas creates a more stable but less growing audience with an average satisfaction of about 75 points (Sahu et al., 2023). Big data technology represents the speed at which we process big data, and it also means that we need comparable computing and storage technologies. Hadoop is an open-source framework specifically designed to handle big data problems. In addition to the problems caused by a large amount of data, data transmission is also an important part of big data research (Anshari et al., 2023). These big data machines represent a huge challenge for communications networks, which need to be able to transfer data in greater volume and at higher speeds. In addition, Moore's Law provides insight into the rapid advances being made in technology: the number of transistors that can fit onto a silicon chip doubles every 18 to 24 months, meaning memory storage grows exponentially. As data is growing exponentially, some other key issues include.

Table 6 provides a comprehensive overview of the performance indicators of the SVM model in predicting audience satisfaction for the web drama, serving as a foundation for the quantitative evaluation and comparison of the model's effectiveness. The analysis necessitates the processing of substantial datasets, which requires more sophisticated methods than traditional statistical techniques. Common approaches in big data analysis employed in this context include cluster analysis, genetic algorithms, natural language processing, machine learning, predictive modeling, regression models, social network analysis, signal processing, and data visualization (Xie et al., 2023). Big data processing and special analysis can help to discard the traditional static treatment process and march towards a dynamic process. Indeed, the methodology of making timely decisions through the analysis of information data had already replaced the traditional processing method, and big data processing and analysis has gradually replaced traditional analysis method in modern business, war, national policy decision, and so on (Xu et al., 2020). Big data influence

Audience Satisfaction Prediction of Chinese Web Drama Based on Big Data Technology and Support Vector Machine

recommendation systems, personalized suggestion systems, face recognition, and payment methods deeply in our life and society (Xu et al., 2020). We believe that the big data utilization in companies and research institutions is about to profoundly affect their decisions and outcomes.

Table 6: Model Predictive Performance Evaluation

Evaluation Indicators	Describe	Numerical Value
Accuracy	Proportion of the number of samples correctly predicted to the total number of samples.	85%+
Recall Rate	The proportion of the number of samples that are actually positive and predicted to be positive to the number of samples that are actually positive.	82%+
F1 Value	Harmonic averages of accuracy and recall for comprehensive evaluation of model performance.	83%+

4. Results and Discussion

With increased influence, especially that of suspense dramas, the consumption landscape of entertainment has changed today. Digital platforms are further extending, and knowing the audience preference is very important for the producers or stakeholders. The present study will delve into the predictability of big data and support vector machines to bring out the changing dynamics between audience satisfaction and drama content. These results are supported by wider research that suggests genre, production quality, and actor performance are key drivers in viewer satisfaction. While online suspense dramas are popular because of their complicated plots and rich emotions, other genres such as comedy and costume dramas also have high satisfaction rates. This represents a diversified demand from audiences and thus requires great attention from producers on various tastes. The model adopted in this study provides a firm grounding not only for the complexities of the narrative but also forms a serious base for prediction of audience responses in an aggressive, competitive industry.

For the study of audience satisfaction with web dramas, a series of measures have been taken to guarantee the security of communication from the audience and the protection of their privacy. The purpose is to meet regulatory and ethical requirements standards, but also to gain confidence and credibility among the audience in respect of the research process and results. Efforts will be continuously made to revise these measures and make them perfect so that the participants may provide a research environment which is far more safe and reliable. Advanced prediction models, such as SVM, are important to be implemented to increase the accuracy of audience satisfaction forecasts, thus assisting data-driven decisions over entertainment content production. It is also critical to consider the changing taste of the audience since the quality of the plot, acting, and production level determines the level of satisfaction. These findings point out the constant need to update prediction models, increase data, and use new machine learning techniques since market trends change quite suddenly. Meanwhile, research on audience analysis should keep a balance between using very sophisticated technologies and paying much attention to issues of ethical concern such as data privacy and security.

These findings of the study have great implications for understanding what determines audience satisfaction in Chinese web dramas. The prediction, with quite high accuracy, was made using big data technology supported by an SVM model. The analysis indicates that factors such as quality of the plot, actor performance, and production values exert a considerable effect on audience perception. Comedy and costume dramas receive high satisfaction scores, revealing that audiences have diversified tastes for entertainment with genres that ensure emotional resonance and aesthetic appeal. Moreover, the findings show that actor popularity serves as a strong driver of engagement, showing a very critical linkage between social media presence and viewer rating. By pinpointing the key drivers of satisfaction in this study, the technical potential of big data in the entertainment sector is underlined, yet at the same time, practical ways for web drama producers to work on increasing the appeal of their content are also provided.

The present study, based on big data technology and support vector machine methodologies, presents a comprehensive analysis to predict audience satisfaction about Chinese web dramas. Massive data in respect of viewership of web drama and user evaluations are collected, analysed, and a robust predictive model for audience satisfaction was constructed. The accuracy and reliability of this model were effectively validated in rigorous testing (Rosati et al., 2023). This predictive model not only boasts high predictive accuracy but also identifies the pivotal factors that influence audience satisfaction, offering targeted optimization recommendations to web drama producers (Wang, 2024). Based on the comparison analysis of different web dramas in terms of audience satisfaction, it could be concluded that plot quality, actor performance, and production standards are major variables affecting viewer satisfaction. These findings have important implications for web drama producers in that they show where to improve the elements which create a more satisfying experience (Yang, 2023). Concluding, this study provides a firm technical support for enhancing audience satisfaction prediction for Chinese network dramas and a good decision framework. We will continually optimize the forecast model in the future and increase the source of data for the maximum effect to innovate the development of web drama. Besides, we would like to look upon the insights derived from this research as an important guide and reference to audience studies in the other cultural sectors.

From the figure 5, it can be observed that a positive relationship exists between the popularity indices of actors, which are comprehensively evaluated by the number of followers on social media and the frequency of online searches, and click-through numbers on the two platforms iQiyi and Tencent. Indexing above 1,100 received more than 50 million hits; that was an almost 30% increase from those who were in the bracket between 1,200 and 1,300. On the other hand, actors who registered less than 1,000 in terms of popularity, their averages seriously slumped under 10 million hits. The actor's popularity hence influential to audience engagement. Furthermore, from the different visualization of audience viewing behaviour, it can also be noticed that the longer the viewing time and the higher the broadcast frequency, the higher the satisfaction of the audience, while more pauses and skips result in negative relations with satisfaction (Zhu et al., 2022). It reflects that an engaging plot and good performance by actors create more satisfaction for an audience, while the dissatisfied viewer would always think about pausing and skipping the content.

Audience Satisfaction Prediction of Chinese Web Drama Based on Big Data Technology and Support Vector Machine

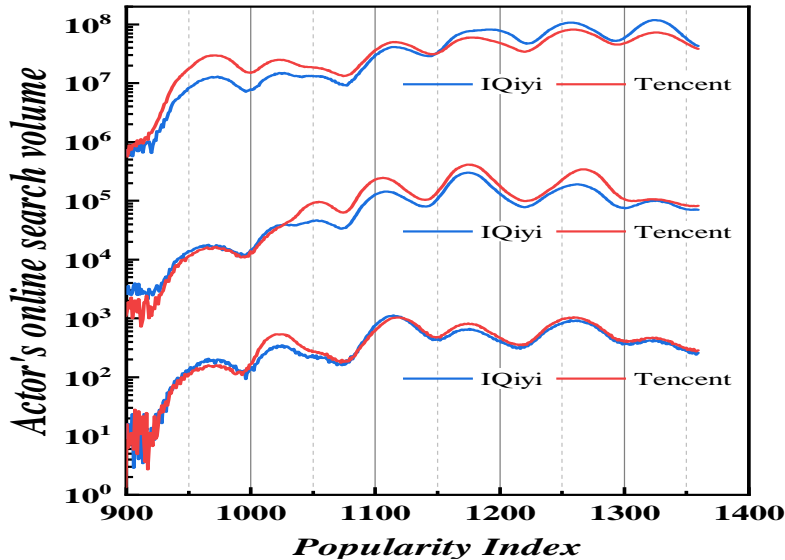


Figure 5: Relationship Chart between Actor Popularity and Satisfaction

The discussion is thinned down to the comparison of genres for web dramas. Comical, while costume drama, enjoys stable satisfaction rates; the thriller and suspense genres are more variable. This might be explained by the fact that these genres are niche themselves, with specific viewers who can be more critical or have higher expectations toward plot coherence and character development. Yet again, the data are suggestive that though thrillers and suspense dramas might give one a very loyal audience, erratic satisfaction measures certainly pose a question to their competent retention. In fact, producers of such genres would do better if they focus on enhancing their narrative complexity and refining plot twists to engage an increasingly demanding audience. While big data technology can be integrated with SVM as a strategic advantage in the interpretation and forecast of audience satisfaction regarding web dramas, the findings indicate that it is about continuous adaptation to the changing trend of viewer preferences, especially in the case of suspense drama, whose plot development may play a key role. It is recommended that producers and platforms use such insights in refining their content and sticking closer to audience expectations. This research further adds to the contemporary debate in the merging of entertainment with technology, shedding light on how data-driven approaches have immense potential in content creation. While it is exciting to see how this growing industry will evolve, it will be even more interesting to see if, in the future, such a predictive model could increasingly be put to use in refining creative output and, more importantly, audience satisfaction.

5. Conclusion

The study adopted big data technology and the model Support Vector Machine in testing audience satisfaction with Chinese web drama, focusing on the high correlation between actors' popularity and viewership. Data from platforms such as iQIYI and Tencent showed that when an actor's popularity index exceeded 1100, large audiences would be attracted, and the views would burst into more than 50 million in those

dramas. As the popularity index arose to 1200-1300, the percentage in viewership rate reached 30% or higher, thereby proving that actors' popularity is in positive relation to drama performance. On the contrary, less than 10 million views are received by those actors whose popularity index was below 1000, thereby pointing out that, in fact, the popularity of actors is what viewers consider. Based on the findings from the research, some development strategies include increasing actors' interaction with the audience and appearing to targeted viewer groups for increased satisfaction and performance. In all, the research findings prove the predictive power of big data and the SVM model, providing lessons that might be helpful to enhance the web drama market with increasing applications of big data.

6. Limitations

Despite these considerable strengths of this test, a few limitations need to be pointed out. First, the study was predominantly based on data regarding audience satisfaction from major platforms only, which cannot fully represent viewership from lesser-known or international platforms. In addition, the research did not consider the cultural and regional preferences that may influence drama reception differently across diverse demographics. Besides, the satisfaction of audiences is naturally subjective and may depend on many personal and cultural factors, which are hardly quantifiable or able to be included into the model. This suggests a natural limitation in the prediction of satisfaction across all kinds of viewers. Last but not least, even though the SVM model gave reliable predictions, further research should be done to embed more machine learning algorithms in order to compare their efficiency in audience satisfaction predictions.

7. Future Directions

In the future, the scope of this study might be made wider by covering more diversified sources of data, which include smaller platforms and international viewership. This could fetch greater knowledge of the consumption patterns of web dramas across the world. Future researchers might investigate how real-time social media interactions related to the same web drama affect audience satisfaction, given that very often social media can take a critical role in shaping audience perceptions of such content. Moreover, sentiment analysis tools, which measure audience emotional responses in real time, can fine-tune the predictions. Knowing just how specific elements of the content-suspense, humour-capture the emotional states of the viewer may yield more fine-grained insights into satisfaction drivers. Further tuning of other machine learning models, including deep learning algorithms, can also provide further refined predictive capabilities and yield a better understanding of the long-term trends in viewer preferences. Moreover, the impact this might have on audience satisfaction through real-time social media interactions introduces a more dynamic layer to predictive models. Expansion of data on international platforms and regional variations may further lead to a better view of the global pattern of consumption. Also, deep learning algorithms could be engaged together with SVM to further enhance the accuracy of satisfaction predictions, opening up new prospects for personalized content creation.

8. Implications

Implications of this study are great and significant to web drama producers, operators of the platform, and marketers. It would be specific with which ways content developers can have their production tuning meet the expectations of the audience, especially in genres of variable levels of satisfaction. Platform operators could use the predictive model developed in the present study to improve the recommendation algorithms in order to increase user engagement through recommending content aligned with the preference of the viewer. With this clarity brought in regarding the popularity of the actors, marketers easily get a way to provide more focused and fruitful promotional campaigns to increase interest and satisfaction among viewers.

8.1 Theoretical Implications

The current paper contributes significantly to the theoretical understanding of audience satisfaction in digital entertainment contexts by employing big data technology and introducing the SVM method for the estimation of viewer responses, thereby expanding the theory scope of "use and satisfaction". This paper also offers empirical support for the view that theories on audience behaviour remain relevant for digital contexts, especially when explaining how plot complexity and genre preference modulate satisfaction. Beyond that, this paper enriches the discourse on media content distribution while giving new perspectives on how data-driven models will be able to influence academic study in the realm of media consumption patterns.

8.2 Practical Implications

The research is an actual guide for producers in regard to the main determinants of audience satisfaction: plot and actor performance. This will, in turn, help develop decisions with respect to content creation, more specifically suspense dramas, by showing the level of innovation in the narration and quality of the production. This predictive model done in this research can also be used by the video platforms in enhancing their recommendation algorithms to provide users with personalized content recommendations that increase user engagement. Furthermore, strong linkage in actor popularity and viewer ratings suggests that marketing strategies should rest on leveraging social media presence to build a larger audience and viewership.

References

- Akbar, M., Ahmad, I., Mirza, M., Ali, M., & Barmavatu, P. (2024). Enhanced authentication for de-duplication of big data on cloud storage system using machine learning approach. *Cluster Computing*, 27(3), 3683-3702. <https://doi.org/10.1007/s10586-023-04171-y>
- Alemdar, K. D., & Çodur, M. Y. (2024). A New Approach to Detect Driver Distraction to Ensure Traffic Safety and Prevent Traffic Accidents: Image Processing and MCDM. *Sustainability*, 16(17), 7642. <https://doi.org/10.3390/su16177642>
- Alsayat, A., & Ahmadi, H. (2023). A hybrid method using ensembles of neural network and text mining for learner satisfaction analysis from big datasets in online learning platform. *Neural Processing Letters*, 55(3), 3267-3303. <https://doi.org/10.1007/s11063-022-11009-y>

- Anshari, M., Syafrudin, M., Tan, A., Fitriyani, N. L., & Alas, Y. (2023). Optimisation of knowledge management (KM) with machine learning (ML) Enabled. *Information*, 14(1), 35. <https://doi.org/10.3390/info14010035>
- Aziz, F. A., Irfangi, A., & Dharin, A. (2022). The Role of Information Technology and Supply Chain Collaboration in Enhancing Innovation. *International Journal Of Construction Supply Chain Management*, 12(2), 143-156. <https://doi.org/10.14424/ijcscm120222-143-156>
- Basyouni, A., Abdelkader, H., Elkilani, W. S., Alharbi, A., Xiao, Y., & Ali, A. H. (2024). A Suicidal Ideation Detection Framework on Social Media using Machine Learning and Genetic Algorithms. *IEEE Access*. <https://doi.org/10.1109/ACCESS.2024.3454796>
- Chen, B., Fan, Y., Lan, W., Liu, J., Cao, C., & Gao, Y. (2022). Fuzzy support vector machine with graph for classifying imbalanced datasets. *Neurocomputing*, 514, 296-312. <https://doi.org/10.1016/j.neucom.2022.09.139>
- Dudzik, W., Nalepa, J., & Kawulok, M. (2024). Ensembles of evolutionarily-constructed support vector machine cascades. *Knowledge-Based Systems*, 288, 111490. <https://doi.org/10.1016/j.knosys.2024.111490>
- Gordillo-Rodriguez, M. T., Pineda, A., & Gómez, J. D. F. (2023). Brand Community and Symbolic Interactionism: A Literature Review. *Review of Communication Research*, 11, 1-32. <https://doi.org/10.5680/RCR.V11.1>
- Gutta, L. M., Bammidi, T. R., Batchu, R. K., & Kanchepu, N. (2024). Real-Time Revelations: Advanced Data Analysis Techniques. *International Journal of Sustainable Development Through AI, ML and IoT*, 3(1), 1-22. <https://ijsdai.com/index.php/IJSDAI/article/view/33>
- Jang, H. E., Kim, S. H., Jeon, J. S., & Oh, J. H. (2023). Visual attributes of thumbnails in predicting youtube brand channel views in the marketing digitalization era. *IEEE Transactions on Computational Social Systems*. 2329-924X. <https://doi.org/10.1109/TCSS.2023.3289410>
- Jingmiao, Y. (2024). Visual Communication Design for Mobile Learning Apps: User Interface Usability and Learning Engagement. *Comunicar: Revista Científica de Comunicación y Educación*(78), 234-248. <https://doi.org/10.58262/V32I78.19>
- Khetarpal, P., Nagpal, N., Alhelou, H. H., Siano, P., & Al-Numay, M. (2024). Noisy and non-stationary power quality disturbance classification based on adaptive segmentation empirical wavelet transform and support vector machine. *Computers and Electrical Engineering*, 118, 109346. <https://doi.org/10.1016/j.compeleceng.2024.109346>
- Li, X., Zhang, D., Zheng, Y., Hong, W., Wang, W., Xia, J., & Lv, Z. (2023). Evolutionary computation-based machine learning for smart city high-dimensional big data analytics. *Applied Soft Computing*, 133, 109955. <https://doi.org/10.1016/j.asoc.2022.109955>
- Liang, Z., & Zhang, L. (2022). Uncertainty-aware twin support vector machines. *Pattern Recognition*, 129, 108706. <https://doi.org/10.1016/j.patcog.2022.108706>
- Mathur, S., Chaba, Y., & Noliya, A. (2023). Performance analysis of support vector machine learning based carrier aggregation resource scheduling in 5g mobile communication. *Procedia Computer Science*, 218, 2776-2785. <https://doi.org/10.1016/j.procs.2023.01.249>
- Mi, C., Li, M., & Wulandari, A. F. (2022). Predicting video views of web series based on comment sentiment analysis and improved stacking ensemble model.

Audience Satisfaction Prediction of Chinese Web Drama Based on Big Data Technology and Support Vector Machine

- Electronic Commerce Research*, 1-28. <https://doi.org/10.1007/s10660-022-09642-9>
- Moshawrab, M., Adda, M., Bouzouane, A., Ibrahim, H., & Raad, A. (2023). PolyFLAG_SVM: a Polymorphic Federated Learning Aggregation of Gradients Support Vector Machines Framework. *Procedia Computer Science*, 224, 139-146. <https://doi.org/10.1016/j.procs.2023.09.021>
- Nieto-Ferrando, J., Gómez-Morales, B., & Castro-Mariño, D. (2023). Audiovisual Fiction, Tourism, and Audience Studies: A Literature Review. *Review of Communication Research*, 11, 88-126. <https://doi.org/10.5680/RCR.V11.4>
- Ponmalar, A., & Dhanakoti, V. (2022). An intrusion detection approach using ensemble support vector machine based chaos game optimization algorithm in big data platform. *Applied Soft Computing*, 116, 108295. <https://doi.org/10.1016/j.asoc.2021.108295>
- Razzak, I., Zafar, K., Imran, M., & Xu, G. (2020). Randomized nonlinear one-class support vector machines with bounded loss function to detect of outliers for large scale IoT data. *Future Generation Computer Systems*, 112, 715-723. <https://doi.org/10.1016/j.future.2020.05.045>
- Rosati, R., Romeo, L., Cecchini, G., Tonetto, F., Viti, P., Mancini, A., & Frontoni, E. (2023). From knowledge-based to big data analytic model: a novel IoT and machine learning based decision support system for predictive maintenance in Industry 4.0. *Journal of Intelligent Manufacturing*, 34(1), 107-121. <https://doi.org/10.1007/s10845-022-01960-x>
- Sahu, S., Kumar, R., Long, H. V., & Shafi, P. M. (2023). Early-production stage prediction of movies success using K-fold hybrid deep ensemble learning model. *Multimedia Tools and Applications*, 82(3), 4031-4061. <https://doi.org/10.1007/s11042-022-13448-0>
- Selmy, H. A., Mohamed, H. K., & Medhat, W. (2023). Big data analytics deep learning techniques and applications: A survey. *Information Systems*, 102318. <https://doi.org/10.1016/j.is.2023.102318>
- Shardaghly, T. H. (2024). A Pragma Stylistic Analysis of Aggression in Hillary Clinton's Speech on Trump. *Eurasian Journal of Applied Linguistics*, 10(1), 269-278. <http://dx.doi.org/10.32601/ejal.10123>
- Shmueli, G., & Polak, J. (2024). Practical time series forecasting with r: A hands-on guide. Axelrod schnell publishers. <https://search.worldcat.org/title/957532009>
- Simonov, A., Ursu, R. M., & Zheng, C. (2023). Suspense and surprise in media product design: Evidence from twitch. *Journal of Marketing Research*, 60(1), 1-24. <https://doi.org/10.1177/00222437221108653>
- Wang, T. C. (2024). Deep learning-based prediction and revenue optimization for online platform user journeys. *Quant. Financ. Econ*, 1, 1-28. <http://dx.doi.org/10.3934/QFE.2024001>
- Xie, X., Sun, F., Qian, J., Guo, L., Zhang, R., Ye, X., & Wang, Z. (2023). Laplacian Lp norm least squares twin support vector machine. *Pattern Recognition*, 136, 109192. <https://doi.org/10.1016/j.patcog.2022.109192>
- Xu, J., Tan, W., & Li, T. (2020). Predicting fan blade icing by using particle swarm optimization and support vector machine algorithm. *Computers & Electrical Engineering*, 87, 106751. <https://doi.org/10.1016/j.compeleceng.2020.106751>

- Yang, S. (2023). Big Tech, Market Boom and Product Quality: Evidence from Drama-Making in China. *Market Boom and Product Quality: Evidence from Drama-Making in China*. <https://dx.doi.org/10.2139/ssrn.4392683>
- Zhao, X., Sun, B., Wu, N., Zeng, R., Geng, R., & He, Z. (2024). A new short-term wind power prediction methodology based on linear and nonlinear hybrid models. *Computers & Industrial Engineering*, 196, 110477. <https://doi.org/10.1016/j.cie.2024.110477>
- Zheng, Y., Ge, Y., Muhsen, S., Wang, S., Elkamchouchi, D. H., Ali, E., & Ali, H. E. (2023). New ridge regression, artificial neural networks and support vector machine for wind speed prediction. *Advances in Engineering Software*, 179, 103426. <https://doi.org/10.1016/j.advengsoft.2023.103426>
- Zhu, W., Song, Y., & Xiao, Y. (2022). Huberized one-class support vector machine with truncated loss function in the primal space. *Advances in Engineering Software*, 173, 103208. <https://doi.org/10.1016/j.advengsoft.2022.103208>