Herding Behavior-Driven Comedy Film Rating Convergence: A Study of the Moderating Effects of Emotional Intensity

Yuyang Wang
Beijing Jiaotong University, No.3, Shangyuan Village, Haidian District, Beijing, China

Keywords: Movie Ratings, Herding Behavior, Sentiment Analysis.

Abstract: The study examines the herding effect in comedy films, particularly how user ratings and comments influence rating behavior. Using data from the Douban movie platform, sentiment and time-series analyses were employed to assess rating convergence and the amplification effect of emotional intensity. Key findings include: low-scored films experience large initial score variance fluctuations that stabilize over time, with emotional factors initially worsening fluctuations but promoting long-term consistency; medium-rated films' emotional resonance varies by subject, significantly affecting convergence speed; and high-scored films show obvious rating convergence, less influenced by emotional intensity. Audience rating behavior is impacted by time, film quality, emotional intensity, and social interaction. The study offers a new angle for the film industry's marketing strategies, advising rating platforms to enhance sentiment analysis algorithms for more accurate emotional tendency capture and improved scoring system reliability.

1 INTRODUCTION

With the widespread adoption of the Internet and social media, user-generated content (UGC) has emerged as a crucial medium for individuals to express their opinions. Among various forms of UGC, online reviews are particularly prevalent across industries such as film, dining, and e-commerce (Chevalier and Mayzlin, 2006). For the film industry, as a kind of online word-of-mouth, users' online comments not only affect other audiences' movie-watching choices, but also may have a profound impact on film box office and long-term word-of-mouth (Samrat et al., 2024). Especially on platforms such as Douban, user ratings and comments have become important criteria to measure the quality of movies.

However, user rating behavior can be significantly influenced by pre-existing average ratings, leading to convergence phenomena, wherein users tend to align their ratings more closely with the existing consensus. The behavioral tendency, termed the herding effect, reflects how individuals adjust their evaluations in response to the prevailing ratings. Furthermore, the sentiment embedded in prior reviews can trigger emotional resonance among users, thereby influencing their own rating and evaluation behavior (Qu et al., 2024). The emotional contagion effect may further exacerbate the herding effect in online rating environments. Additionally, different types of movies have different types of audiences (Urszula, 2019), which may also lead to different degrees of herding effect among the audience. For example, the degree of herding effect of comedy movies may be more dependent on the emotional tendency of users because of its relaxed and pleasant characteristics. The questions are of great significance for understanding the psychological mechanism behind user evaluation behavior and the optimization of marketing strategies in the film industry.

2 LITERATURE REVIEW

The herding effect refers to the tendency of individuals to conform to group behaviors under social influence, often disregarding their independent judgment. The phenomenon has been extensively studied across disciplines such as social psychology, economics, and behavioral finance. In social psychology, herding is closely related to conformity, where individuals adopt group behaviors to gain social acceptance or avoid exclusion, even when those behaviors contradict personal opinions (Bond

Wang, Y.
Herding Behavior-Driven Comedy Film Rating Convergence: A Study of the Moderating Effects of Emotional Intensity.
DOI: 10.5220/0013841700004719

In *Proceedings of the 2nd International Conference on E-commerce and Modern Logistics (ICEML 2025)*, pages 207-213
ISBN: 978-989-758-775-7

and Smith, 1996). Salganik et al. (2006), through their music download experiment, revealed how social influences and herding can lead to inequality and unpredictability in cultural products. The study found that when participants were presented with information about the choices of others, their choice behavior was significantly affected, resulting in the popularity of certain cultural products far higher than others. The phenomenon is also relatively common in movie rating, and users' rating behavior may be guided by the existing average rating, and then show the phenomenon of convergence. Hao et al. (2019), taking the film industry as an entry point, studied the decision-making behaviors of consumers under the influence of social media and found that the herd effect of consumers still exists in the film field. Lee et al. (2015) compared the effects of strangers and friends on user-generated movie reviews and found that users would be influenced by the mainstream ratings of previous movies, and would follow the mainstream or deliberately express different opinions instead of their own opinions.

Sentiment analysis is the automatic identification and extraction of emotional tendencies in text through computational methods. Common methods include dictionary-based methods and machine learning methods. Pang and Lee (2008) conducted a systematic review of opinion mining and sentiment classification, highlighting that dictionary-based methods demonstrate superior accuracy in domains such as film criticism. Taboada et al. (2011) further explored the efficacy of various sentiment dictionaries, emphasizing optimization strategies to enhance classification precision.

Furthermore, audience evaluation behaviors differ across movie genres. Urszula (2019) noted that genre-specific factors influence rating patterns. Hao Yuan et al. (2020) believe that audience's evaluation of comedy movies is often influenced by factors such as plot, actor performance and visual effects, so the rating of comedy movies is more dependent on users' emotional tendency (Zhang et al., 2021).

To sum up, existing studies have revealed the widespread existence of herding effect in fields including movie ratings. However, for the specific genre of comedy movies, the herding mechanism of user ratings still needs further research. In particular, existing research has focused less on how the intensity of reviews affects the herding effect of ratings, and whether the effect differs across films of different ratings levels. Through empirical analysis, the study will explore the role of user rating and comment emotion in the herd effect of comedy movies, and provide a new perspective for the research in the field.

3 RESEARCH PROBLEM

The aim of the study is to investigate the influence of user ratings and review sentiment on herding effects in comedy films. Specifically, the research will address the following key questions:

1. In comedy movies, is there a convergence of user ratings, that is, does the degree of deviation from the average rating decrease with the change of the average rating?
2. Does the emotional intensity of reviews amplify the herding effect of ratings?
3. Does the herding effect differ among films with different ratings?

4 DATA PROCESSING

4.1 Data Source

The dataset utilized in the study was sourced from the Douban Movie Platform, with data collection conducted via the Octopus web crawler platform. The raw dataset included various attributes such as movie titles, user ratings, short review texts, and release dates. To ensure data representativeness and validity, the study adopted a series of rigorous selection and processing measures. First of all, as one of the largest film communities in China, Douban Film platform has a huge user base and rich movie data, and its ratings and reviews have high representativeness and reference value. Users participate in evaluation and scoring spontaneously on the Douban platform, and the data can better restore the average views of the general public (Liu, 2018). In addition, Douban's film ratings and reviews are often cited and referenced by media reports and even professional film reviews, further demonstrating the authority and reliability of its data. (Shi, 2024)

Secondly, the Octopus web crawler platform was employed due to its user-friendly, code-free operation and robust data extraction capabilities. It offers built-in templates, extensive network data capture support, and advanced deduplication and filtering functions, ensuring the quality and uniqueness of the extracted data while minimizing redundancy. The features enhance the efficiency and accuracy of data acquisition, thereby strengthening data reliability.In the process of data collection, in order to control the consistency of film types, the study focuses on comedy films (Douban film label includes comedy), and randomly selects films released in recent years as research objects. According to the overall film score of the platform, movies are divided into low grades with less than 6 points, medium grades with 6-9 points and high grades with more than 9 points. The low and middle grades selected five films each, while the high grades selected all three films due to the relatively small number of films. 400 reviews and ratings were extracted for each film according to its popularity, which ensured that each sample film had enough data for analysis, and covered reviews of different popularity to improve the comprehensiveness and representation of the data. In the data preprocessing stage, blank comments and unrated invalid comments were cleared, and 4982 valid comments and their scores were obtained. Through such data screening and cleaning, the quality and effectiveness of the data are further guaranteed, and a reliable data basis is provided for subsequent research and analysis.

4.2 Data Measurement

In the present study, emotion dictionaries are utilized to determine the emotional intensity of reviews. As the general Chinese emotion dictionaries are not capable of explaining movie reviews, emotion dictionaries suitable for movie reviews are used in the emotion analysis (Wang, et al., 2022), which not only merged and cleaned the well-known Hownet dictionaries and NTUSD dictionaries. A dedicated sentiment dictionary has also been constructed specifically for the field of film, which is able to more accurately capture the emotional tendencies in film reviews. Compared to the improved Hownet dictionary and NTUSD dictionary, the dictionary is 8.1% and 10% more accurate at capturing emotions, respectively. In the dictionary, the weight of positive emotion words is set to 1, and the weight of negative emotion words is -1. At the same time, the dictionary of degree adverbs is used to match degree adverbs in reviews and assign corresponding weights according to their impact on emotional intensity. In the calculation of emotion value, the weight of emotion words, degree adverbs and negative words is comprehensively considered, according to the Wang（2023）’s formula:

$$\begin{aligned}&\text{Affective value}\\&=\left|\sum\big((\text{positive affective word weight}\\&-\text{negative affective word weight})\\&\times\text{degree adverb weight}\big)\right|\end{aligned} \quad (1)$$

Effectively quantifies the emotional tendency and intensity of the comments.

In addition, to facilitate the calculation of the emotion-weighted score, the max-minimum normalization is used to map the emotive value to the interval [0,1] to eliminate the impact of dimension:

$$x' = \frac{X - x_{min}}{x_{max} - x_{min}} \quad (2)$$

Then, by multiplying the mapped star rating and the normalized emotion intensity, a composite index "emotion weighted score" is generated to reflect both the objective value of the user's rating and the subjective intensity of the emotion tendency, so as to analyze whether the emotion of the comments can promote the herd effect (Yao et al.,2017). For example, a review with a five-star rating 5 combined with strong positive emotion 2.0 has a weighted score of 9.

In time series analysis, the dynamic average score of a movie is an important index to measure the trend of movie rating over time. However, the movie history score data of Douban film platform is not open. In order to analyse the trend of the score over time, all data is calculated in "days". The cumulative average of all star ratings and emotion-weighted ratings from the date of the film's earliest appearance (i.e. the film's release date) is calculated separately from the date of release to the date of release, which can simulate the dynamic change of movie ratings over time during the film's release.

According to the cumulative average score sum of each day, the variance of the corresponding day can be calculated, which can reflect the convergence degree of the audience for the movie score and the change of the convergence degree after emotional weighting.

However, the time series fluctuation caused by the number of film reviews with heat attenuation leads to sharp fluctuations in the variance in the later period. The study proposes an improved smoothing strategy based on linear weighted moving average (WMA) (Box et al.,2015). By introducing the standardized index mechanism of review participation, the model can keep the sensitivity of recent data while effectively reducing the variance oscillation caused by the decrease of the number of comments in the later period.

Specifically, first build a standardized metric based on comment engagement:

$$NormCom_i = \frac{C_i}{\max\limits_{k \in \Omega}(C_k)} \tag{3}$$

Among them:

C_i represents the number of comments in the i th time unit (day)

Ω is the collection of time units covered by the sliding window

The numerator and denominator standardization constrained the index range to [0,1], reflecting the relative intensity of comment activity on that day

A standardized metric of comment engagement is then used as a standardized weight calculation:

$$\text{Smoothing rating/} \\ \text{emotion weighted rating variance} = \frac{\sum_{i=0}^{W}(\text{the variance} \times NormCom_i)}{\sum_{i=0}^{W}(NormCom_i)} \tag{4}$$

In the study, W=5 is selected as the window size

The smoothed score/emotional-weighted score variance is obtained, which ultimately reduces the violent vibration caused by late fluctuations, and can better capture the stability of scoring trends and changes in emotional consensus.

5 REGRESSION ANALYSIS

To examine the temporal variation in rating dispersion and emotional intensity, an exponential regression model is employed to analyze herding effects in rating convergence and the amplification effect of sentiment-driven herding. Regression model $y(t) = a \cdot e^{-bt} + c$ takes time t as the independent variable and score variance or emotion weighted score variance as the dependent variable, where:

1. a is the magnitude of the initial variance,
2. b represents the rate of decay,
3. c represents the level of variance that is stable over time.

The regression coefficient a represents the initial variance of the movie score. The sign and magnitude of regression coefficient b reflect the direction and rate of volatility change. If the regression coefficient is greater than zero, it indicates that the variance has a decreasing trend over time, that is, the degree of dispersion of scores or emotions has decreased, which can be regarded as the embodiment of herding effect. Furthermore, by comparing the regression coefficients of smoothing score variance and affective weighted score variance, we can further judge whether affective factors aggravate the convergence phenomenon. Among them, when processing the regression coefficient of emotion weighted score variance, the previous emotion weighted score was calculated by multiplying star rating and normalized emotion intensity, which may lead to compression of regression coefficient. Therefore, before processing emotion weighted score variance, divide each sample point by the normalized average emotion intensity of the movie. The effect of the compression of regression coefficient caused by multiplication on the measurement of herding effect should be minimized. Next, the regression results will be analysed horizontally from three aspects: low score, middle score and high grade.

5.1 Low-Grade Films

Table 5.1: Regression results of low-grade films.

Movie Title	Rating Variance Regression	Emotional Rating Variance Regression
Life on the Road	y=1.1807e^(-0.3603x) +2.2916, R²=0.7189	y=2.5716e^(-0.2421x) +1.8081, R²=0.7935
Ex-lover 4: Early Marriage	y=0.2832e^(-0.0144x) +2.4910, R²=0.7910	y=0.7366e^(-0.0149x) +2.3812, R²=0.7440
Detective Chinatown 3	y=0.0609e^(0.0698x) +1.9713, R²=0.7156	y=-0.1444e^(3.1805x) +1.8864, R²=0.6962
The Extraordinary Journey of the Mozart from Outer Space	y=0.6940e^(-0.5473x) +1.8281, R²=0.9192	y=0.5514e^(-0.3082x) +1.6098, R²=0.7748
Super Family	y=0.0519e^(2.9342x) +1.6344, R²=0.0246	y=0.0686e^(0.0058x) +1.3023, R²=0.1855

The regression results of low-grade films show that their score variance has the characteristics of high initial fluctuation and rapid decay. Taking *Life on the Road* as an example, the initial value of the score variance a+c=1.1807+2.2916=3.4723, and the decay rate b=0.3603, indicating that controversial content such as plot logic holes caused sharp differences in audience evaluation in the early stage. However, with the passage of time, the score rapidly converges to the stable value. The phenomenon may be due to the "dispute-driven discussion" mechanism of low-scoring films: Extremely bad reviews attract more viewers to participate in the review, but as the discussion progresses, some viewers re-examine the work, resulting in a rapid decline in the variance. For example, the initial value of *The Extraordinary Journey of the Mozart from Outer Space* is 2.5221, and the decay rate b=0.5473, which further verifies the dynamic pattern of "high opening and low walking" in low-score films.

However, the low-grade films' decay rate of affective score variance is generally lower than that of ordinary score. For example, the decay rate of the emotion score of *Life on the Road* is lower than that of the ordinary score, and the asymptotic value c is lower. The suggests that although affective weighting failed to accelerate convergence, it stabilized the score at a more concentrated range by filtering out extreme emotional noise. The result is consistent with the theory of emotion regulation in psychology. Sentiment analysis indirectly inhibits the long-term proliferation of controversial comments by identifying negative emotional intensity.

5.2 Mid-Grade Films

Table 2: Regression results of mid-grade films.

Movie Title	**Rating Variance Regression**	**Emotional Rating Variance Regression**
Grabbing the Doll	y=0.1709e^(0.0944x) +2.9905, R^2=0.5988	y=1.3646e^(0.1316x) +2.6414, R^2=0.8213
Manjianghong	y=0.8444e^(-0.0745x) +2.9220, R^2=0.8374	y=1.3284e^(-0.2077x) +2.5380, R^2=0.9558
Hot and Spicy	y=0.5118e^(-0.0520x) +3.2106, R^2=0.9364	y=0.9267e^(-0.0446x) +3.3420, R^2=0.9285
Article 20	y=0.2223e^(0.4021x) +3.3506, R^2=0.7082	y=0.6803e^(0.1822x) +3.0684, R^2=0.6676
Flying High 2	y=0.2471e^(0.2362x) +3.4188, R^2=0.3440	y=1.8370e^(0.2009x) +2.9015, R^2=0.9400

The rating variance of mid-grade films shows significant decay rate differentiation. Taking "Hot and Hot" as an example, the decay rate of ordinary score variance b=0.0520, asymptotic value c=3.2106, while the decay rate of emotional score variance b=0.0446, asymptotic value c=3.3420, and R^2 increased from 0.6864 to 0.9285. The indicates that although the emotional intensity does not accelerate the decline, by directing the audience to focus on the core theme of female growth, the impact of minor disputes on the rating is reduced, and the explanatory power of the model is significantly improved.

In contrast, in Article 20, the decay rate of ordinary score variance b=0.4021, asymptotic value c=3.3506, decay rate of affective score variance b=0.1822, asymptotic value c=3.0684, and R^2 decreased from 0.7082 to 0.6676. The shows that sentiment analysis fails to reconcile the group antagonism caused by legal disputes, reflecting the failure of the simplified assumption of the exponential model for complex social issues.

It is worth noting that the variance decay rate of emotion score b=0.2077 is significantly faster than that of ordinary score (b=0.0745), the asymptotic value c drops from 2.9220 to 2.5380, and R^2 jumps from 0.8374 to 0.9558. The result may be due to the audience's collective memory resonance of historical themes: emotion weighting accelerates the convergence of opinions by capturing positive emotions such as "family feelings" and "heroism". The phenomenon is called emotional mobilization effect in communication studies, that is, specific emotional labels can strengthen group identification and suppress differences (Cui Junli et al, 2024).

5.3 High-Grade Films

Table 3: Regression results of high-grade films

Movie Title	**Rating Variance Regression**	**Emotional Rating Variance Regression**

Coco	y=1.2127e^(-0.0452x) +1.5340, R²=0.8901	y=0.7074e^(0.4166x) +1.3301, R²=0.0436
Dying to Survive	y=0.5601e^(0.1576x) +1.6231, R²=0.7002	y=2.0864e^(0.2195x) +1.5855, R²=0.5470
Zootopia	y=1.0759e^(-0.0727x) +1.7293, R²=0.7554	y=1.6746e^(-0.0480x) +1.3786, R²=0.5968

The score variance of high grade films is characterized by low initial fluctuation and slow decay. Taking Zootopia as an example, the initial value of common score variance is 2.8052, decay rate b=0.0727, asymptotic value c=1.7293, R^2=0.7554; The initial value of emotion score variance a+c=1.6746+1.3786= 3.0532, decay rate b=0.0480, asymptotic value c=1.3786, R^2=0.5968. The indicates that the audience's initial evaluation of the classic works is highly converging, and the emotional intensity fails to significantly change its dynamics. The quantization bias of affective weight reduces the explanatory power of the model.

However, the variance regression of emotion score in Coco is not significant at all, which may be due to its artistic expression of the theme of "life and death cycle", which exceeds the quantitative range of traditional emotion dictionaries, leading to the failure of the model. In addition, the initial variance value of emotion score 3.6719 of Failing to Survive is significantly higher than the initial value of common score 2.1832, reflecting the infiltration of external discussion of social issues into the score, forming a hypercinematic evaluation dynamic (Lin Hongtong, 2020).

6 CONCLUSION

The study employs an exponential regression model to uncover the intricate mechanisms governing the dynamic rating behavior of comedy films on Douban. The findings indicate that audience rating convergence is influenced not only by the passage of time but also by factors such as film quality classification, emotional intensity of reviews, and the depth of audience engagement. In the case of low-rated films, the presence of controversial content often results in pronounced fluctuations in initial ratings. While sentiment analysis helps mitigate the short-term impact of extreme reviews, it remains insufficient in reversing the non-linear trend of word-of-mouth decline over time. Although emotional resonance can accelerate rating convergence for certain films, it fails to alleviate group polarization caused by controversial themes or social issues. Conversely, for high-rated films, the initial evaluations tend to be highly consistent, leading to greater rating stability. However, the quantitative bias introduced by sentiment-weighted scores limits the explanatory power of deep emotional interactions in the cases. The innovative contribution of the study is as follows: First, it focuses on the specific type of comedy film, and deeply discusses the herding mechanism of user rating, which fills the gap of existing research in the field; Second, it not only focuses on the convergence of user ratings, but also comprehensively considers the impact of the emotional strength of reviews on the herd effect, as well as the differences of the effect in films with different ratings, providing a more detailed and comprehensive perspective for the marketing strategy of the film industry.

Despite the contributions, the study has several limitations. First, sentiment analysis algorithms still have room for improvement. Due to the implicit nature and ambiguity of Chinese language expressions, reliance on static sentiment dictionaries alone may fail to fully capture context-dependent and implicit emotions in reviews. Future research could enhance accuracy by incorporating irony detection and contextual semantic analysis techniques. Second, as the dataset is sourced exclusively from Douban, audience feedback from other social media platforms is not considered, which may limit the generalizability of the findings. Future studies could validate and extend the conclusions by integrating data from multiple online platforms. Based on the findings and deficiencies, the following coping strategies are suggested for movie scoring platforms and producers: Movie scoring platforms should continue to optimize sentiment analysis algorithms and introduce irony detection and contextual semantic analysis technologies to more accurately capture users' emotional tendencies. For example, natural language processing technology is used to analyse comments deeply and identify complex emotional expressions such as irony and metaphor, so as to improve the accuracy and reliability of emotion analysis. At the same time, the platform needs to build a real-time scoring monitoring mechanism to dynamically monitor movies of different scoring grades and timely warn of abnormal fluctuations. Through data analysis and machine learning algorithms, the system can automatically identify abnormal patterns in scores to provide decision support for platform managers.

Producers need to develop differentiated strategies for different grades based on film ratings and critical data. Low-grade films should respond quickly to the controversial evaluation and adjust the marketing strategy or film content in time to recover the reputation; Mid-range films can strengthen emotional labels, guide audiences to form a consensus, and improve the stability of word-of-mouth; High grade films should make use of in-depth reviews and fan activities to maintain and consolidate their good reputation. In addition, producers can work with scoring platforms to obtain sentiment analysis data from audience reviews to gain insight into audience emotional needs and preferences, so that targeted optimization can be made during film production. For example, according to the analysis of the audience's emotional tendency towards a comedy film, the plot setting, actor performance and visual effects of the film are adjusted to better meet the audience's expectations. By implementing the strategies, film rating platforms and producers can better navigate the challenges posed by the herding effect, enhance the reliability of film rating systems, refine industry marketing strategies, and ultimately deliver higher-quality content and services to audiences.

REFERENCES

Bond, R., Smith, P., 1996. *Culture and conformity: A meta-analysis of studies using Asch's (1952b, 1956) line judgment task. Psychological Bulletin, 119(1), 111-137.*

Box, G., Jenkins, G., Reinsel, G., Ljung, G., 2015. *Time Series Analysis: Forecasting and Control (5th ed.). Wiley.*

Chevalier, J., Mayzlin, D., 2006. *The effect of word of mouth on sales: Online book reviews. Journal of Marketing Research, 43(3), 345-354.*

Cui, J., Jiang, H., 2024. *Emotional Mobilization and Resistance Practices in Hot Events on Douyin Platform—A Case Study of the Phenomenon of Public Copywriting Relay Communication. Progress in psychology, 14 (9), 508-515. https://doi.org/10.12677/ap.2024.149675*

Gupta, S., et al., 2024. *How consumers evaluate movies on online platforms? Investigating the role of consumer engagement and external engagement. Journal of Business Research, 176, 114613. https://doi.org/10.1016/j.jbusres.2024.114613*

Hao, Y., Zou, P., Li, Y., et al., 2009. *Empirical study on the impact of online review sentiment on sales revenue based on movie panel data. Management Review, 21(10), 120-128.*

Hao, X., Chen, X., 2019. *The Herding Effect of Consumption Decision of Experience Products and Mechanism Explanation -An Empirical Interpretation of Consumer Behavior in Film Industry[J]. Chinese Journal of Management Science, 2019, 27(11): 176-188. 10.16381/j.cnki.issn1003-207x.2019.11.018*

Lee, Y. J., Hosanagar, K., Tan, Y., 2015. *Do I Follow My Friends or the Crowd? Information Cascades in Online Movie Ratings. Management Science, 61(9), 2241-2258. https://doi.org/10.1287/mnsc.2014.2082*

Lin, H., 2020. *Film Narrative Techniques and Methods: Analysis of Classic Films. China Film Press.*

Liu, Y., 2018. *Analysis of the Current Situation and Public Opinion of Film Rating Communities at Home and Abroad[J]. Audio-visual, 2018(1): 34-35.*

Liu, W., Ma, Z, Wei, X., 2019. *A meta-analysis of the impact of crowding on consumer emotions and shopping reactions. Acta Psychologica Sinica, 52(1), 16-28.*

Pang, B., Lee, L., 2008. *Opinion mining and sentiment analysis. Foundations and Trends in Information Retrieval, 2(1-2), 1-135.*

Qu, W., et al., 2024. *A study on the influencing factors of movie consumer satisfaction based on text mining. Sino-American Arts and Sciences, 13(4), 115. https://doi.org/10.12677/sa.2024.134115.*

Salganik, M., Dodds, P., Watts, D., 2006. *Experimental study of inequality and unpredictability in an artificial cultural market. Science, 311(5762), 854-856.*

Shi T., 2024. *Analysis of LDA Theme Model for Douban Movie Reviews Based on Text Mining —Taking the Movie "Let the Bullets Fly" as an Example. Journalism and Communications Journalism and communication science, 2024, 12(1), 23-28. https://doi.org/10.12677/JC.2024.121004*

Świerczyńska-Kaczor, U., 2019. *Exploring the relationship between viewer experience and movie genre – A study based on text mining of online movie reviews. Problemy Zarządzania – Management Issues, 17(5), 85-9. https://doi.org/10.7172/1644-9584.85.9*

Taboada, M., Brooke, J., Tofiloski, M., Ide, N., Szatrowski, T., 2011. *Lexicon-based methods for sentiment analysis. Computational Linguistics, 37(2), 267-307.*

Wang, T., Zhang, Z., 2022. *Research on the construction method of sentiment dictionary for movie reviews. Computer and Digital Engineering, 50(4), 843-848.*

Wang, w., Wu, j., 2023. *Sentiment Analysis of microblog text based on Sentiment Dictionary and Semantic Rule Set [J]. Computer Science and Applications, 2023, 13(4): 754-763. DOI: 10.12677/csa.2023.134074*

Yao, Y., 2017. *Empirical study on sentiment analysis of online reviews based on user rating preferences. (Dissertation). Shanghai: Tongji University.*

Zhang, Z., Qiang, C., Duan, S., 2021. *A review of personalized movie recommendation algorithms. Computer Knowledge and Technology, 17(22), 1-10*