Proceedings of CONF-CDS 2025 Symposium: Data Visualization Methods for Evaluation
DOI: 10.54254/2755-2721/2025.PO24512

Research on the Dissemination of Chinese Animated Culture from the Perspective of TikTok Comments: A Case Study of Nezha: The Devil Child Roars into the Sea

Yingxi Wei

Anhui University, Hefei, China
weiyingxi0112@gmail.com

Abstract. As the global prominence of Chinese cultural artifacts continues to expand, the investigation of their cross-cultural reception has emerged as a pivotal area of scholarly inquiry. This study examines comments made by overseas TikTok users on the animated film Nezha: The Devil Child Roars into the Sea. Utilizing the DrissionPage instrument, a corpus of over 4,200 English comments was amassed. Subsequent to data acquisition, preprocessing was executed employing NLTK, followed by manual sentiment annotation facilitated through Label Studio. A bert-base-uncased model was then trained for sentiment classification, achieving a validation accuracy of 89.01%. Results indicate that overseas audiences generally responded positively to the film, particularly appreciating its animation quality, mythological themes, and character development. However, cultural discount was also observed, including feelings of cultural unfamiliarity, difficulty understanding the plot, and criticism of cultural adaptations. Based on these findings, strategies are proposed to mitigate cultural discount effects, such as supplementing cultural background information, adopting localized narrative approaches, strengthening cultural IP branding, and collaborating with local content creators. This research offers empirical evidence and actionable recommendations for the global distribution of Chinese animation and cultural content.

Keywords: Cross-Cultural Communication, Sentiment Analysis, TikTok Comments, Nezha, BERT Model

1. Introduction

In the context of globalization, cultural exchange has become a crucial avenue for developing national cultures. The diversification of media formats has rendered film and television essential conduits for cross-cultural communication, serving as significant bridges between disparate cultures [1]. In recent years, China's film industry has not only achieved remarkable success domestically but also garnered significant popularity overseas with films such as the Nezha series, The Battle at Lake Changjin, Wolf Warrior 2, and Hi, Mom. However, the international release of these films often faces challenges in cultural dissemination, such as the attenuation of cultural value due to foreign audiences' limited comprehension of Chinese culture or broader issues of cultural

© 2025 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2025.PO24512

incompatibility [2]. This phenomenon, known as “cultural discount,” refers to the reduction in meaning interpretation or value perception of cultural products in cross-cultural reception, caused by differences in cultural backgrounds, aesthetic preferences, and value systems [3].

Given that the film Nezha: The Devil Child Roars into the Sea (hereafter Nezha 2) achieved notable box office success upon its release in early 2025 and has generated substantial and recent discussion on TikTok, it holds significant research value. This study collects and analyzes overseas TikTok users’ comments on the film using Python-based text analysis techniques. The aim is to analyze the international perception of Chinese culture and audience predilections for its constituents, thereby furnishing actionable methodologies to mitigate cultural discount and foster its global propagation.

This study utilizes social media analytics to empirically evaluate cross-cultural reception of Chinese cultural exports, formulating frameworks to quantify cultural discount via textual analysis. It employs Natural Language Processing (NLP) and sentiment analysis of User-Generated Content(UGC) to broaden cultural diffusion research methodologies.

2. Literature review

Regarding the theme of Sino-foreign cultural integration or the global dissemination of Chinese culture, Liu employed Python-driven data processing to compile a database of reader comments on English translations of The Analects, examining translation quality, ideological perception, aesthetics, and China's cultural image to improve translation quality and cultural exchange [4]. Li et al. used Python web scraping to collect overseas player comments on Genshin Impact from YouTube. Utilizing WordCloud, Gensim, and the Wenxin model, they examined players' preferences for traditional Chinese cultural elements in the game, facilitating global dissemination of Chinese cultural heritage [5]. Zou Shuangji used Black Myth: Wukong as a case study, collecting comments under its official YouTube trailer through Python-based scraping. Employing ROST software for analytical purposes, the investigation scrutinized the feedback of international players through cognitive, affective, and behavioral lenses, thereby offering insights for the global dissemination of Chinese-developed games and augmenting their cross-cultural communication and cultural impact [6]. Chen conducted sentiment analysis on literary and other textual materials to trace the developmental trajectory of Western culture. The study proposed a BERT-BiLSTM-based method for short text tasks, outperforming traditional deep learning models to achieve more accurate sentiment analysis [7].

In summary, current academic research on this topic predominantly focuses on games and literary works as research objects, with Python and ROST software frequently employed as analytical tools. This investigation consequently elects to concentrate on a comparatively underexamined medium—cinema and animation—as its focal point of inquiry, utilizing Python-based tools for data collection and analysis to mitigate cultural discount in cross-cultural transmission and promote the more effective global dissemination of Chinese culture.

3. Research design

This study follows a structured methodology comprising the following steps:

(1)Using DrissionPage: one of Python third-party libraries, comments related to the film Nezha 2 were collected from the TikTok platform. Preprocessing entailed lemmatization, case normalization, stop word elimination, and emoji filtering, yielding standardized comments. These were tokenized for word frequency analysis and word cloud visualization.

DOI: 10.54254/2755-2721/2025.PO24512

(2)Label Studio was employed for manual sentiment annotation of the collected comments. Subsequently, comments categorized as positive or negative were used to train and test the bert-base-uncased model.

4. Related theories

4.1. DrissionPage tool

DrissionPage is a Python-based web automation tool that integrates the strengths of Selenium and Requests. DrissionPage integrates Driver and Session modes, facilitating browser control and network packet management. By repurposing the Chromium protocol and Requests library, it streamlines transitions between automation and data extraction, minimizing development intricacy and code bulk.

4.2. BERT model

BERT (Bidirectional Encoder Representations from Transformers) is a deep learning-based language representation model proposed by Google AI in October 2018, whose core architecture is based on the Transformer encoder. The Transformer model, introduced for Google's translation systems in 2017, consists of an encoder and decoder. BERT utilizes solely the encoder, available in BERT-base and BERT-large versions with 12 and 24 encoder layers, respectively. Compared to traditional models such as recurrent neural networks (RNNs) and long short-term memory networks (LSTMs), the Transformer architecture exhibits superior text modeling capabilities and better utilization of high-performance hardware like GPUs, thereby enabling large-scale data training [8]. BERT's advancements have enabled it to attain state-of-the-art performance in diverse NLP tasks, representing a pivotal milestone in language model evolution.

4.3. Sentiment analysis in social media contexts

Sentiment analysis, also known as opinion mining, is a subfield of NLP that focuses on automatically identifying the emotional tone expressed in written texts. The exponential rise of social media data and increased access to public opinions have rendered sentiment analysis indispensable for gauging public sentiment across diverse domains [9].

5. Experimental analysis

5.1. Data collection

DrissionPage, integrating Selenium's automated browsing with Requests' efficient handling, was employed to extract comment data from TikTok, adeptly navigating dynamic web structures. The tool successfully loaded and extracted comments through keyword searches and simulated page scrolling. After removing duplicate and non-English comments, 4,506 comment texts were obtained and saved in xlsx format, providing the corpus for subsequent analysis.

5.2. Data preprocessing

Before starting the experiments, a series of preprocessing steps were conducted on the comment data, including deduplication, noise reduction, tokenization, removal of stop words and punctuation,

DOI: 10.54254/2755-2721/2025.PO24512

and lemmatization to ensure standardized data quality and improved analytical accuracy. Initially, a for-loop was used to identify and remove duplicate comments. Emojis embedded in the original data were cleared through Python's re (regular expression) library by creating pattern objects with re.compile() and replacing matches with empty strings via re.sub(). After these steps, 4,207 valid comments were retained.

To ensure that personal names such as "ne zha" and "ao bing" would not be split during tokenization, the study applied regular expression substitution with re.sub(), replacing "ne zha" with "ne_zha" and "ao bing" with "ao_bing". Then, the word_tokenize() function from the NLTK library was used for tokenization, and all words were converted to lowercase with lower(). Subsequently, the pos_tag() function from NLTK was used for part-of-speech tagging, followed by lemmatization using the lemmatize() function. Finally, meaningless or low-value words, such as "i," "it," "the," "watch," and "look," were included in a custom stopword list and removed alongside punctuation marks.

5.3. Word frequency statistics and word cloud analysis

To better understand the focal points of overseas audiences regarding the film, word frequency statistics were first conducted to calculate the occurrence frequency of each word in the comments. During this process, personal names were standardized by unifying "ne_zha" and "ao_bing" into "nezha" and "aobing," respectively. Word frequency statistics allowed for the precise extraction of high-frequency terms, revealing major themes and core elements discussed by audiences.

A word cloud, created using Python's WordCloud library, visualizes text data by word frequency, providing an intuitive summary of overseas audiences' sentiment and focus on Nezha 2 (Figure 1).

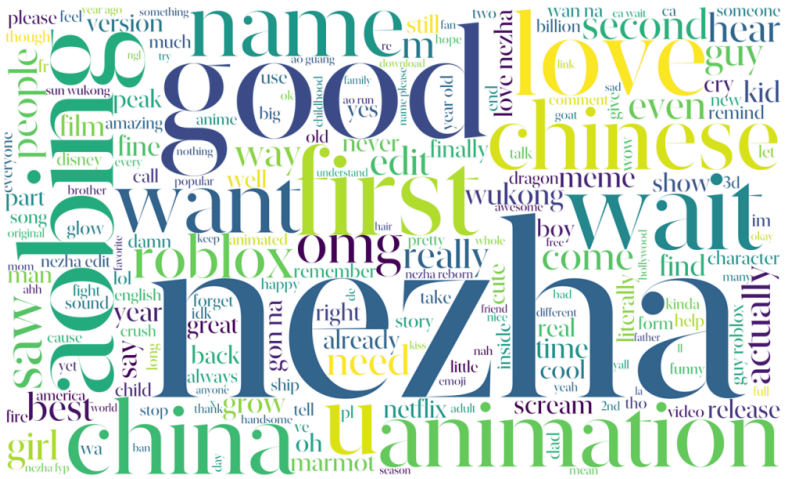

Figure 1: Word cloud

From the word cloud, it can be observed that overseas audiences generally expressed positive and enthusiastic attitudes toward the film. Words such as "good," "love," "great," and "peak" appear prominently, indicating high appreciation and enjoyment. These frequent terms suggest that the film received favorable evaluations, especially regarding its animation quality, storytelling, and character development. Additionally, the prominence of words like "China" and "Chinese" highlights audience interest and recognition of the film's Chinese cultural elements, such as its traditional mythology and artistic style. The word "animation" appears prominently and reflects a strong interest in the technical quality of the animation. Specific character names such as "nezha" and "ao

DOI: 10.54254/2755-2721/2025.PO24512

bing" occupy significant space in the word cloud, indicating intense audience engagement with these figures.

5.4. Sentiment analysis

Using the Label Studio platform, 4,206 comments were manually annotated for sentiment. After excluding neutral comments, 1,122 remained, with a positive-to-negative ratio of approximately 9:1. Twenty percent of the annotated data was used as a training set to fine-tune the bert-base-uncased model. In contrast, the remaining data served as the validation set. The model was trained with a batch size 16, a learning rate 2e-5, and 10 epochs. The final validation accuracy reached 0.8901, indicating a relatively high classification performance. Among the 992 positive comments collected, it is evident that overseas audiences expressed generally favorable sentiments toward Nezha 2.

Many comments reflected a strong interest in Chinese mythology, as evidenced by statements like "I love Chinese mythology," and "Same, I love it." This suggests that the film sparked curiosity and appreciation for traditional Chinese culture. In addition, the character designs garnered considerable attention. For instance, comments such as "That white dragon boy is hot" illustrate how visually appealing characters can enhance emotional engagement among viewers. Beyond the film itself, some audiences expressed admiration for Chinese cultural products more broadly. Remarks like "I love Chinese people, man. They make good games and good movies" indicate that high-quality cultural exports can contribute positively to national image and cultural soft power. Collectively, these findings elucidate that the articulation of traditional culture via engaging narrative techniques and emotionally evocative imagery can markedly enhance audience engagement and facilitate cultural assimilation on an international level.

Negative comments from overseas viewers can be broadly categorized into four types: unfamiliarity, cross-cultural comparison, narrative confusion, and criticism of mythological adaptation. Many comments reveal unfamiliarity with the film, as seen in expressions like "Never seen or heard of either film" and "Ne Zha 2? I didn't even know there was a 1 Imaooo." The expressions "never heard" and "didn't know" imply a significant lack of awareness among international audiences regarding the film's existence, indicative of the series' constrained international recognition attributable to inadequate global marketing strategies and cultural contextualization.

Another typical pattern involves cross-cultural comparisons. For example, comments such as "I don't know why people so obsessed with that animated movie… there is no way that movie is better than Toy Story 3 or Up or Grave of the Fireflies or Spirited Away" show that some viewers measure Ne Zha 2 against globally acclaimed animations and tend to favor the latter. It reflects the persistent impact of the cultural discount is evident in its role in modulating aesthetic predilections.

Some viewers also reported difficulty understanding the film, as evidenced by remarks like "I watched this movie and was so confused" and "I genuinely don't understand the movie at all." These answers indicate narrative uncertainty, which most likely results from an unfamiliarity with the cultural and mythological context necessary to understand the story and character development.

Additionally, several comments question the accuracy of mythological representations. For instance, "She's so pretty, but she's originally a man in Chinese mythology, so it isn't that accurate, considering this movie is more of a parody or depiction of Chinese mythology (take Sun Wukong for example)" reflects concerns about the liberties taken in character design. Words like "originally a man," "not accurate," and "strange" suggest that while some viewers possess prior knowledge of Chinese mythology, they expect cultural adaptations to strike a balance between innovation and authenticity.

DOI: 10.54254/2755-2721/2025.PO24512

6. Discussion and suggestions

6.1. Discussion

Analysis of over 4,200 English comments from TikTok reveals that Nezha 2 generally received positive feedback from overseas audiences.

The word cloud analysis frequently appears to have highly positive words such as “good,” “love,” “great,” and “amazing.” At the same time, the sentiment classification results indicate that positive comments account for 88% of the total. This implies that the film's visual presentation, character growth, and cultural components effectively captivated the attention and emotional resonance of viewers worldwide, highlighting the potent potential of Chinese animated productions in international communication.

However, signs of cultural discount also emerged beyond the positive feedback, reflecting barriers and cognitive gaps in the cultural export process. Specifically, some viewers expressed unfamiliarity with the film, indicating that Chinese animation’s international visibility remains limited. Some comments revealed a preference for familiar Western cultural products through direct comparisons to movies like Toy Story and Spirited Away, suggesting that cultural backgrounds still profoundly influence aesthetic judgments. Numerous viewers expressed confusion over the plot and character dynamics, underscoring the challenges of comprehending Chinese mythological contexts absent prior cultural familiarity. Additionally, critiques of Nezha's gender representation indicate audience awareness of the tension between cultural innovation and traditional authenticity.

These results suggest that cultural discount is a complicated interaction between cultural unfamiliarity, inadequate communication techniques, and disparities in storytelling customs rather than a straightforward manifestation of rejection. Therefore, although cultural discount is still an inevitable occurrence in cross-cultural communication, it may be methodically reduced by using focused approaches in audience interaction, contextual adaptation, and content production.

6.2. Strategies for enhancing the global dissemination of Chinese culture

Based on the word frequency and sentiment analysis findings of over 4,200 English comments on TikTok, Nezha 2 was generally well-received abroad, particularly in terms of animation quality, character development, and cultural richness. However, signs of cultural discount also surfaced. Drawing from the four identified issues, several strategic recommendations can be made to optimize future cultural exports.

First and foremost, it is essential to add cultural background knowledge. Character profiles, behind-the-scenes featurettes, or brief cultural explanations should be included in international adaptations of works that have rich traditional Chinese aspects. By lowering cultural barriers and encouraging emotional involvement, these additions can assist global audiences in developing a fundamental knowledge.

In addition, narrative strategies should strive for a balance between internationally recognizable storytelling techniques and the authenticity of local cultural expression. While maintaining the core of Chinese cultural identity, adopting narrative structures and character-building methods familiar to global audiences may enhance cross-cultural accessibility and aesthetic resonance. This approach echoes the theoretical framework of “hybridity,” which highlights the dynamic interplay between global forms and local meanings [10].

Another key strategy involves strengthening the branding and serialization of Chinese cultural intellectual properties. As demonstrated in Jin’s analysis of the Korean Wave, systematic IP

DOI: 10.54254/2755-2721/2025.PO24512

management and the strategic use of short video platforms can contribute to forming globally recognizable cultural symbol systems [11]. Chinese creators might investigate the establishment of sustainable character universes, exemplified by a "Nezha Universe," utilizing short video marketing, merchandising, and transmedia storytelling to cultivate enduring audience engagement.

Moreover, the role of localized dissemination on social media deserves greater attention. Local content creators function not only as distributors but also as cultural interpreters who help bridge cultural gaps. Collaborating with influential local figures on platforms like TikTok and YouTube allows for reinterpretation, explanation, and creative engagement with Chinese cultural products in ways that resonate more deeply with international audiences.

Equally important is the need to balance cultural innovation with respect for tradition. As Venuti posits, cultural communicators must navigate between “visible innovation” and “invisible fidelity” [13]. In order to prevent any misunderstandings that could impede cultural acceptability, character and narrative adaptations should incorporate contemporary aesthetic features while remaining sensitive to audience expectations for cultural authenticity.

7. Conclusion

This study focused on overseas TikTok users’ comments on the animated film Nezha 2, utilizing Python-based data processing and natural language processing techniques to systematically analyze word frequency distribution, sentiment orientation, and cultural acceptance. Data was collected using DrissionPage, while NLTK was employed for data preprocessing and word frequency analysis. Sentiment categories were manually annotated with Label Studio, and a BERT model was trained and validated.

The study indicates that international audiences generally held favorable views of the film, especially its animation and Chinese cultural aspects. However, cultural discount was apparent, marked by unfamiliarity with Chinese mythology, comparisons to known foreign works, and concerns over character reinterpretation. These findings affirm the pertinence of cultural discount theory in cross-cultural communication, underscoring cognitive gaps and communication challenges for Chinese cultural products globally.

This study's limitations include its reliance on TikTok, limiting sample diversity and geographic scope. Future research should consider multi-platform data integration (e.g., YouTube, Instagram) to enhance representativeness and incorporate multimodal features (images, audio, emojis) for comprehensive sentiment and semantic analysis [14].

References

[1] Yuan, X. J., & Zhu, C. Q. (2025). Perceptual degree, communicative context, and heterogeneous components: An examination of the "cultural discount" phenomenon in film and television works in cross-cultural communication. Journal of Changsha University of Science and Technology (Social Science Edition) , 40 (02), 113-121.

[2] Li, X. (2025). Factors influencing cross-cultural communication of Chinese fantasy films from the cultural discount perspective. International Communication of Chinese Culture, 12 , 77–92.

[3] Hoskins, C., & Mirus, R. (1988). Reasons for the U.S. dominance of the international trade in television programmes. Media, Culture & Society .

[4] Liu, H. (2020, December). A study on the reader reception of English translations of The Analects-data analysis with Python. In 2020 5th International Conference on Mechanical, Control and Computer Engineering (ICMCCE) (pp. 2381-2387). IEEE.

[5] Li, L. D., Tang, W. Q., Liu, S. H., & Lin, R. C. (2024). Preference analysis of overseas game players on Chinese cultural elements based on text mining—Taking "Genshin Impact" as an example. Modern Information Technology , 8 (22), 111-116.

DOI: 10.54254/2755-2721/2025.PO24512

[6] Zou, S. J. (2025). A study on the cross-cultural communication effect of domestic games—Based on the comment analysis of the "Black Myth: Wukong" trailer. Rong Mei , (03), 34-40.
[7] Chen, M. (2022). Emotion analysis based on deep learning with application to research on the development of Western culture. Frontiers in Psychology, 13 , 911686.
[8] Liu, H., Zhang, Z. X., & Wang, Y. F. (2021). A review of main optimization and improvement methods of BERT model. Data Analysis and Knowledge Discovery , 5 (1), 3-15.
[9] Wang, Q. Y., Shi, S. C., & Wang, H. J. (2025). A review of text sentiment analysis. Software Guide , 24 (01), 193-202.
[10] Kraidy, M. (2006). Hybridity, or the cultural logic of globalization . Temple University Press.
[11] Jin, D. Y. (2016). New Korean Wave: Transnational cultural power in the age of social media .
[12] Lobato, R., & Thomas, J. (2015). The informal media economy .
[13] Venuti, L. (1995). The translator’s invisibility .
[14] Kumar, S., & Sebastian, T. M. (2012). Sentiment analysis: A perspective on its past, present and future. International Journal of Intelligent Systems and Applications , 4 (10), 1–14.