Chapter 2

Literature Review

2.1 Overview

This chapter reviews the literature related to this research topic, mainly covering the global response of movies, sentiment analysis methods, and the impact of cultural differences on audience emotional responses. By combing through existing research, this paper clearly points out the shortcomings of current research and further explains the value of this study in theory and practice.

2.2 How movies are rated differently around the world

As Chinese films gradually enter the international market, audiences in different countries have increasingly different opinions on the same film. Studies have found that the narrative style, theme setting, cultural elements, etc. of a film will affect the audience's emotional response. For example, Chinese mythology films such as "Nezha: The Devil Boy Conquers the Dragon King" have been highly accepted in Asia, but have performed mediocrely in countries with different cultural backgrounds.

Despite the increasing international dissemination of Chinese films, research on understanding these differences through quantitative sentiment analysis is still very limited, which has become an important starting point for this article.

2.3 Sentiment Analysis and Its Application in Film Evaluation

Sentiment Analysis is an important branch of Natural Language Processing (NLP), which is mainly used to identify the subjective emotional tendencies expressed in text. By analyzing comments, posts, articles, etc., sentiment analysis can determine whether the user's attitude is positive, negative or neutral. In film review research, sentiment analysis is widely used to understand the audience's emotional response, predict box office, and measure film reputation.
Sentiment analysis is mainly divided into the following methods:

(1) VADER (Valence Aware Dictionary and sEntiment Reasoner)

VADER is a dictionary-based sentiment analysis method designed for short social media texts (such as tweets and movie reviews). It scores the words in a sentence using a built-in sentiment dictionary and combines grammatical structures such as punctuation, capital letters, and degree adverbs to enhance the recognition of sentiment intensity. The advantages of VADER are fast speed, suitability for English comments, and good processing of informal language. The disadvantage is that it is difficult to handle deep semantics or complex contexts.

Applicable scenarios: English short reviews (such as IMDb, Twitter comments), large-scale preliminary classification.

(2) TextBlob

TextBlob is a Python text processing library and a dictionary-based sentiment analysis tool. It uses built-in sentiment dictionaries (such as the Pattern library) to quickly analyze the sentiment polarity and subjectivity of text. TextBlob performs well in the initial classification of English reviews, but is weak in understanding complex expressions or polysemous sentences.

The advantages are: easy to use and high computational efficiency; the disadvantages are: difficult to adapt to multiple languages and weak ability to recognize emotion reversal.

Applicable scenarios: fast classification of English movie reviews and prototype model testing.

(3) Naive Bayes Naive Bayes classifier

Naive Bayes is a classic supervised learning algorithm based on probability theory and used for sentiment classification of texts. It assumes that each word appears independently in the text and determines the sentiment tendency by calculating the probability of words appearing in different sentiment categories. This method has fast training speed and stable results and is often used for medium-scale sentiment text analysis tasks.

The advantage is that it is suitable for multi-category sentiment classification; the disadvantage is that it seriously ignores the contextual relationship between words.

Applicable scenarios: classification of positive and negative emotions in movie reviews, and initial model for processing multilingual reviews.

(4) Support Vector Machine (SVM)

SVM is a supervised machine learning algorithm that is particularly suitable for handling classification problems of high-dimensional text data. It distinguishes different categories by finding the best hyperplane so that the classification result has the largest boundary interval. SVM performs stably in sentiment polarity recognition and is particularly suitable for structured and well-labeled comment data.

The advantages are: high classification accuracy; the disadvantages are: high cost of training large-scale data and high sensitivity to text sparsity.

Applicable scenarios: high-precision sentiment classification of movie review big data and unified modeling of cross-language texts.

(5) BERT (Bidirectional Encoder Representations from Transformers)

BERT is a pre-trained language model proposed by Google that performs well in understanding text context. Compared with traditional methods, BERT can capture the meaning of words from the entire sentence and dynamically adjust the emotional tendency of words according to the context. It is one of the most advanced models in current sentiment analysis. It is particularly suitable for processing movie review texts with complex language structures and delicate emotional expressions.

Advantages: strong semantic understanding ability, support for multiple languages in Chinese and English, and high multi-task adaptability; Disadvantages: requires a lot of computing resources and takes a long time to train.

Applicable scenarios: In-depth analysis of emotional expressions on "plot", "characters", "cultural symbols" and other aspects in reviews, and analysis of Chinese reviews.

(6) Aspect-Based Sentiment Analysis

This method not only identifies the overall sentiment tendency of the text, but also identifies the user's emotions towards a specific aspect, such as the specific attitude towards "plot", "image", "character" and "music". This method can capture the audience's specific concerns more finely and is an important supplement to traditional sentiment analysis.

Advantages: The analysis dimension is more refined and the results are more interpretable; Challenges: It is necessary to label or extract aspect words, and the modeling is more complicated.

Applicable scenarios: Film review research with strong explanatory power, such as analyzing what audiences in different countries praise or criticize.

Main models and methods:

Model	Type	Features	Use Case
VADER	Lexicon-based	Lightweight, suitable for social media	Short English texts like Twitter, IMDb reviews
TextBlob	Rule-based	Fast implementation, supports polarity and subjectivity scoring	Basic English text analysis
Naive Bayes	Machine Learning	Simple and efficient, good for structured text	Multiclass sentiment classification
SVM	Machine Learning	High classification accuracy, suitable for large datasets	High-accuracy sentiment polarity classification
BERT	Deep Learning	Understands context, suitable for multilingual tasks	Emotion detection in multilingual movie reviews
RoBERTa, ERNIE	Pretrained Models	Deeper language understanding	High-accuracy sentiment classification for Chinese reviews

Deep learning methods outperform traditional methods in processing texts rich in emotions and context, such as movie reviews, because they can understand context. In addition, Aspect-Based Sentiment Analysis is also gradually being used to extract users' emotional evaluations of specific aspects such as "plot", "soundtrack", and "roles", which is more helpful in understanding the specific reasons behind the differences.

2.4 The impact of cultural differences on emotional responses

Audiences from different cultural backgrounds will have significantly different acceptance of the same movie. Different countries have differences in collectivism, power distance, uncertainty avoidance, etc., which will affect people's emotional expression and acceptance of values. For example, a film that strongly emphasizes family or sacrifice is more likely to be well received in East Asian culture, which is dominated by collectivism, but may be interpreted as overly exaggerated or difficult to empathize with in Western culture, which is dominated by individualism. Although cultural studies are widely present in the fields of communication and cross-cultural communication, analyzing cultural

differences from audience comments in combination with sentiment analysis technology is still a research gap.

2.5 Social media and user review data

With the development of platforms such as IMDb, Douban, and Rotten Tomatoes, moviegoers can freely express their opinions, and these comment data have become an important resource for studying the emotions of movie audiences. Existing studies have used comment data from social platforms to mine public sentiment and analyze emotional attitudes on issues such as policies, brands, and even war.
However, it is still uncommon to systematically apply these methods to the global evaluation analysis of Chinese films, especially the lack of research on emotional comparison of comments from audiences in multiple countries.

2.6 Research gaps and positioning

According to the above analysis, the existing research has the following shortcomings:

- ✧ Most sentiment analysis studies focus only on a single market (such as domestic audiences) and lack cross-national comparative analysis;
- ✧ There is a lack of research that combines cultural background with sentiment data to explain evaluation differences;
- ✧ There is a lack of international audience acceptance analysis for Chinese local cultural films (such as "Nezha 2");
- ✧ Most sentiment analysis only deals with positive and negative classifications and lacks in-depth exploration of aspect-level sentiment.

Therefore, this study will fill these gaps, combine sentiment analysis with cultural difference theory, conduct a systematic analysis of audience reviews of Nezha 2 in multiple countries, explore the commonalities and differences in the evaluations of Nezha 2 in different countries, and help Chinese films better go global.

2.7 Summary

This chapter reviews the literature on global film evaluation, sentiment analysis methods, and the impact of cultural differences, and clearly points out the shortcomings of current research in terms of methods and applications. This article will take "Nezha 2" as an example and use a variety of sentiment analysis techniques to analyze the differences in audience reactions in different cultural contexts around the world. The research results

will provide data support and strategic reference for the global dissemination of Chinese films.

REFERENCES

Briskilal, J., & Subalalitha, C. N. (n.d.). An ensemble model for classifying idioms and literal texts using BERT and RoBERTa. SRM Institute of Science and Technology, India.

Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. Journal of Machine Learning Research, 13(Feb), 281–305.

Bellman, R. (1961). Adaptive control processes: A guided tour. Princeton University Press.

Tan, C. Y., Ong, H. F., Lim, C. H., Tan, M. S., Ooi, E. H., & Wong, K. (n.d.). Amogel: A multi-omics classification framework using associative graph neural networks with prior knowledge for biomarker identification. Monash University Malaysia.

Hsu, W., Lee, M. L., & Goh, K. G. (n.d.). Image mining in IRIS: Integrated retinal information system. National University of Singapore.

Romero, C., Ventura, S., Vasilyeva, E., & Pechenizkiy, M. (n.d.). Class association rule mining from students' test data. Córdoba University & Eindhoven University of Technology.

Manjarres, A. V., Sandoval, L. G. M., & Suárez, M. J. S. (n.d.). Data mining techniques applied in educational environments: Literature review. Politécnico Grancolombiano & Minuto de Dios, Colombia.

Staneviciene, E., Gudoniene, D., Punys, V., & Kukstys, A. (n.d.). A case study on the data mining-based prediction of students' performance for effective and sustainable e-learning. Kaunas University of Technology & Bergen Kommune.

Muzsnay, A., & Szabó, C. (n.d.). Retrieval practice – A tool to narrow the knowledge gap in learning higher mathematics. University of Debrecen & Eötvös Loránd University.

Ghashout, S., Gdura, Y., & Drawil, N. (n.d.). Early prediction of students' academic performance using artificial neural network: A case study in computer engineering department. University of Tripoli.

Ježdík, P. (n.d.). Centralized diagnostics of electronic and electric equipment in vehicles, injector valves testing method. Czech Technical University of Prague.

Alaíz, C., Barbero, Á. L., Fernández, Á. N., & Dorronsoro, J. R. (n.d.). High wind and energy specific models for global production forecast. Universidad Autónoma de Madrid.

Zhuo, M., Wang, P., & Wang, K.-K. (n.d.). Prediction model building for finished tobacco quality based on MIV-BP neural network. China Tobacco Hubei Industrial Co. Ltd.

Diao, Y., Li, Z., Shan, Z., & Li, F. (n.d.). Filter rod quality inspection algorithm based on multi-scale feature fusing attention mechanism. Kunming University of Science and Technology.

Jonuzi, T., Masaad, S., Lupo, A., Domenech Gomez, J. D., Bienstman, P., & Massar, S. (n.d.). Numerical analysis of a self-calibrating time-spatial interleaving photonic convolutional accelerator. VLC Photonics S.L.

Lin, H. X., Zhang, H., Cai, H., Griffin, P., & Liu, A. Q. (n.d.). A quantum photonic chip for binary classification. Nanyang Technological University.