

Chapter 5

5.1 Conclusion

This study uses the film reviews of Nezha 2 as the research object, and uses sentiment classification methods to conduct an in-depth analysis of the differences in sentiment among audiences in China, Malaysia, and the United States. The research process includes data collection, text preprocessing, feature extraction, model training, and result comparison. Finally, TF-IDF, Logistic Regression, random forest, and XGBoost are used for modeling.

There are obvious differences in the viewing emotions of audiences in different countries:

- Chinese reviews were generally positive, reflecting cultural recognition and support for domestic animation;
- Malaysian comments also showed high levels of positive sentiment, with particular attention paid to visual effects and emotional resonance;
- American reviews were more neutral or negative, which may be related to cultural differences, unfamiliarity with narrative style and mythological background.

In terms of model performance, the XGBoost model performed best in terms of accuracy, precision, and F1 score, indicating that deep pre-training models have obvious advantages in capturing text details and contextual semantics.

Overall, sentiment analysis not only reveals the different reactions of audiences across regions to the same film, but also provides data support and inspiration for international film and television communication.

5.2 Future work

Future research can be expanded in the following directions:

1. Introduce real geographic location information: obtain real user regions through platform region tags or IP information to improve the accuracy of regional analysis;
2. Introduction of multilingual and local language models: Adopt regional pre-trained models such as Chinese-BERT and MalayBERT to avoid translation bias;
3. Emotional evolution analysis in the time dimension: Track the emotional changes before and after the movie is released and analyze the emotional trends;