ISSN: 2424-8460(O)
2251-2608(P)

Original Research Article

A comparative study of neural network - based machine translation and human translation: A case study of public display language translation in yunnan tourist attractions

Xinying Xi

Yunnan Minzu University, Kunming, Yunnan, 650031, China

Abstract: With the advancement of the "Belt and Road"initiative, the internationalization degree of Yunnan's tourism is continuously increasing, with 5.8 million inbound tourists received in 2019. This study takes the public display language of Yunnan's tourist attractions as the research object and systematically compares the performance of neural machine translation (NMT) and human translation through a combination of quantitative analysis and qualitative assessment. The study reveals the limitations of machine translation in processing culturally sensitive texts and proposes an "intelligent + human"collaborative translation model along with algorithm optimization suggestions to provide practical references for the field of cultural tourism translation.

Keywords: neural network; machine translation; human translation; translation quality assessment

1. Introduction

As a key medium for cross-cultural communication, the translation quality of public display language in tourist attractions directly affects the travel experience of international tourists and the dissemination effect of national culture. Yunnan, as a multi-ethnic settlement area, contains rich ethnic cultural elements in its scenic area public display language, and this kind of text puts forward high requirements for the cultural adaptability of translation.

2. Research methods and corpus construction of neural machine translation and human translation based on neural networks

2.1. Steps of corpus construction

The study adopted a stratified sampling method and collected 1,000 public display languages in five 4A-level and above scenic spots, including Lijiang Ancient Town, to build a bilingual parallel corpus. The specific steps included: first, determining the sampling framework according to the type and functional zoning of the scenic spot to ensure that the corpus covers different scenarios; second, obtaining the original public display language through on-site shooting, collection of official texts, and other means; and finally, cleaning, transcribing, and completing the bilingual alignment and term annotation of the corpus. To ensure the reliability of the data, the research team conducted multiple rounds of verification on the corpus, eliminating duplicate or ambiguous texts, and ultimately formed a structurally complete research corpus.

2.2. Classification and annotation of the corpus

The corpus was divided into four categories according to its function: directional indication (300 items, 30%), scenic spot description (400 items, 40%), safety warning (200 items, 20%), and service facilities (100 items, 10%). During the classification process, the research team combined the practical function and cultural

ISSN: 2424-8460(O)
2251-2608(P)

connotation of the public display language in tourism to develop a unified annotation standard: for directional indication materials, the accuracy of spatial information was emphasized in the annotation; for scenic spot description materials, the annotation of cultural symbols was emphasized; for safety warning materials, the annotation of the appropriateness of the imperative tone was focused on; and for service facility materials, the consistency of terminology was emphasized. In addition, the corpus introduced a unique ethnic culture annotation system, laying the foundation for subsequent cultural adaptability analysis.

3. Comparison of neural machine translation and human translation

3.1. Accuracy comparison analysis

The experimental results showed that human translation performed outstandingly in the translation of cultural load words (χ^2=32.1.5, $p<0.001$). For example, the machine translation of "Three Courses of Tea"was "Three-course tea"(a literal translation), while the human translation was "Bai Ethnic Three-step Tea"(cultural interpretation), which achieved the dual accurate transmission of semantics and culture by adding ethnic cultural attributes. The study used a contingency table and chi-square test to verify the significant advantage of human translation in dealing with ethnic cultural terms such as "Dongba script"and "Pearl of the Plateau."Machine translation, due to the lack of understanding of cultural context, often has term mistranslation or cultural omission, while human translation ensures the equivalent transmission of the source language's cultural connotation through the reconstruction of the translator's cultural schema.

3.2. Differences in fluency

The independent sample t-test showed that the fluency score of human translation was significantly higher than that of machine translation ($t=18.34$, $p<0.01$), especially in scenic spot description texts, with the largest difference ($\Delta=1.8$ points). Machine translation, limited by the algorithm, often has logical gaps in the processing of long sentence structures and rhetorical conversion, such as translating complex scenic spot introductions into grammatically fragmented expressions; while human translation, based on the translator's language mastery, can adjust sentence patterns according to the target language's expression habits, making the translation more in line with the reading expectations of tourists. Effect size analysis showed that the difference between the two had practical significance (Cohen's d=2.2.7>0.8), indicating the irreplaceability of human translation in improving text readability.

3.3. Comparison of cultural adaptability

Machine translation had systematic defects in the processing of cultural items, with an average cultural adaptability of only 35%, far lower than that of human translation at 85%. Specifically, in directional indication materials, machine translation lacked cultural annotations for the processing of special geographical terms in ethnic areas; in scenic spot description texts, the translation of cultural symbols such as "tie-dye"and "wind, flower, snow, and moon"only stayed at the literal level, failing to convey their ethnic craftsmanship or literary connotations; in safety warning and service facility materials, there were also problems such as the mistranslation of religious taboo words or inaccurate ecological discourse. Human translation, through the cultural mediation of the translator, realized the dual functions of information transmission and cultural dissemination, such as accurately translating "prayer wheel"and supplementing cultural explanations in context.

ISSN: 2424-8460(O)
2251-2608(P)

4. Exploration of human-machine collaborative translation model

In response to the cultural limitations of machine translation, the study proposed an "intelligent + human"collaborative translation model and constructed a quality monitoring model. The model includes five core modules: The cultural sensitivity word detection module, by loading a multi-ethnic sensitive word library and terminology library, can identify taboo words and term mistranslations in the translation in real-time (such as detecting Tibetan taboo words like "kill"or Bai ethnic number taboo "4"); the three-party collaborative quality inspection workflow, based on blockchain technology, realizes multi-dimensional review by language experts, cultural consultants, and local residents, ensuring the transparency and authority of translation decisions; the multi-dimensional quality scoring algorithm dynamically adjusts weights according to different ethnic cultural characteristics (for example, increasing the cultural dimension weight to 0.5 in Tibetan texts), comprehensively evaluating the language, culture, and functional adaptability of the translation; the error feedback mechanism collects error cases (such as the mistranslation of "holy mountain") and retrains the model to promote the continuous evolution of the AI translation system; the system integration main control module integrates various functions to achieve a full-process closed-loop management from detection, review to scoring and optimization. This model not only leverages the efficiency advantages of machine translation but also makes up for the cultural processing deficiencies through human intervention.

5. Application of translation technology

In terms of technical implementation, the collaborative translation system is recommended to adopt a layered architecture: The core engine layer combines rule engine and NLP technology (such as spaCy) to realize cultural detection and uses Scikit-learn to build a quality assessment algorithm; the data layer manages the ethnic terminology library through Neo4j graph database and stores error cases with Elasticsearch; the blockchain layer uses Hyperledger Fabric to realize the notarization of review records; the deployment layer realizes system scalability through a microservice architecture (such as API Gateway, cultural detection microservice, etc.) and realizes asynchronous processing of error feedback and model training through Kafka message queue. In addition, the system also supports multi-dimensional visual quality reports to provide decision-making basis for managers and promote the precise application of translation technology in the field of cultural tourism.

6. Conclusion

This study has confirmed through comparative experiments that human translation still occupies a dominant position in the translation of public display language in Yunnan's tourist attractions, especially in the processing of cultural adaptability. In the future, the coverage of the corpus can be further expanded, and the algorithm research on the translation of ethnic cultural symbols can be deepened to promote the deep integration of translation technology and cultural tourism and help improve the cross-cultural communication efficiency of ethnic regions.

About the author

Xinying Xi (September 2003 -), female, Han ethnicity, from Dali, Yunnan Province. She holds a bachelor's degree and her research direction is Translation Studies. Her postal code is 650504.

ISSN: 2424-8460(O)
2251-2608(P)

References

[1] Lu P ,Xu F . The quality optimization of English–Chinese machine translation based on deep neural networks[J]. Discover Artificial Intelligence, 2025, 5(1): 88-88.

[2] PLOS One Editors. Retraction: The use of machine translation algorithm based on residual and LSTM neural network in translation teaching. [J]. PloS one, 2025, 20(4): e0321329.

[3] Han Q .Research on the Accuracy of Machine Translation in Cross-Cultural Communication Based on Embedded Neural Networks [J]. International Journal of High Speed Electronics and Systems, 2024, (prepublish).

[4] Nouhaila B, Habib A,Abdellah A, et al. Contextualized dynamic meta embeddings based on Gated CNNs and self-attention for Arabic machine translation[J].International Journal of Intelligent Computing and Cybernetics, 2024, 17(3): 605-631.