Journal of http://www.satursonpublishing.com/
Research in Multidisciplinary Methods and Applications
ISSN 2957-3920 (Online) ISSN 3007-7060 (Print)
Volume 4, Issue 10, October 2025

Ethnic Festival LLM Via Prompt and LoRA Fine-Tuning

Wang Dan[1,2], Zhao Weina[1,2,*], Ma Longlong[3], An Bo[4]

1 School of Computer Science, Qinghai Normal University, Xining 810008, China; 2 State Key Laboratory of Tibetan Intelligent Information Processing and Application, Xining 810008, China; 3 Institute of Software, Chinese Academy of Sciences, Beijing 100190, China; 4 Institute of Ethnology and Anthropology, Chinese Academy of Social Sciences, Beijing 100081, China.

*Corresponding to: Zhao Weina

Abstract: As the core carrier of ethnic culture, the construction of a multi-ethnic festival question-and-answer model is conducive to promoting the digital development of ethnic cultural heritage.Addressing the challenges of constructing large language models and the high costs of full fine-tuning, this paper investigates LoRA fine-tuning methods for ethnic festival large language models. It establishes an ethnic festival question-answering model centered on core content from books such as The Chinese Festival Traditional Culture Reader: Collector's Edition and Ethnic Minority Festivals, alongside ethnic festival introductions from Baidu Baike and ethnic websites. First, we constructed datasets and guided the model to generate knowledge question-answer pairs through prompt design. Subsequently, multiple general-purpose large models underwent LoRA fine-tuning. Finally, the optimal fine-tuned model was selected through evaluation and its effectiveness verified. Experimental results show that the fine-tuned ChatGLM4-9B-Chat model achieved improvements of 16.74, 14.25, 13.48, and 16.00 in BLEU-4, ROUGE-1, ROUGE-2, and ROUGE-L metrics respectively compared to the base model. The experiments demonstrate that the selected model can effectively learn and comprehend knowledge related to festival domains, providing accurate answers to user queries. This contributes to promoting the dissemination of ethnic festival culture and advancing the digitalization of festivals.

Keywords: Ethnic festivals, large language models, prompt engineering, LoRA

1 0INTRODUCTION

Ethnic festivals serve as vital bonds sustaining the cultural ecology of multi-ethnic societies. Against the backdrop of rapid digital advancement, developing large-scale ethnic festival question-answering models capable of precisely addressing user needs not only provides the public with efficient and convenient knowledge access channels but also facilitates the interaction, exchange, and integration of ethnic cultures[1].

In recent years, large language models (LLMs) have garnered widespread attention for their exceptional text generation and natural language understanding capabilities[2]. However, their application in specialized domains still faces multiple challenges: on one hand, constructing general-purpose large models requires substantial computational power and data resources, making rapid deployment in niche fields like ethnic culture difficult[3]; on the other hand, models often exhibit "hallucination" when processing domain knowledge—generating content that appears plausible but is factually incorrect[4]. Furthermore, knowledge about ethnic festivals exhibits distinct regional, folkloric, and historical characteristics, areas where existing general-purpose models generally lack deep understanding and precise expression capabilities.

To address these issues, this paper focuses on researching fine-tuning methods for ethnic festival large language models. It aims to construct high-precision ethnic festival question-answering models through low-cost, highly targeted fine-tuning strategies combined with authoritative data sources. The study primarily utilizes classical texts such as the "Reader on Traditional Chinese Festival Culture (Collector's Edition)[5]" and "Ethnic Minority Festivals[6]," alongside authoritative content from Baidu Baike (https://baike. baidu.com/), and the ethnic culture website (http://www.minzu56.net/). Through systematic dataset construction, prompt-guided question-answer pair generation, multi-model LoRA fine-tuning, and performance evaluation, the model achieves precise comprehension and efficient Q&A capabilities regarding ethnic festival knowledge.

2 LARGE LANGUAGE MODEL FOR ETHNIC FESTIVALS BASED ON LORA

Journal of http://www.satursonpublishing.com/
Research in Multidisciplinary Methods and Applications
ISSN 2957-3920 (Online) ISSN 3007-7060 (Print)
Volume 4, Issue 10, October 2025

This paper constructs a multi-ethnic festival knowledge question-answering dataset by generating festival-related questions using prompt engineering large models, based on core content from ethnic festival books and websites. Corresponding answers are then generated by inference-based large models according to the text content. Finally, the constructed Q&A dataset is used to fine-tune the large model through micro-tuning techniques. The fine-tuned model results are evaluated to select the most suitable Q&A model for the ethnic festival domain. The overall framework of this paper is illustrated in Figure 1.

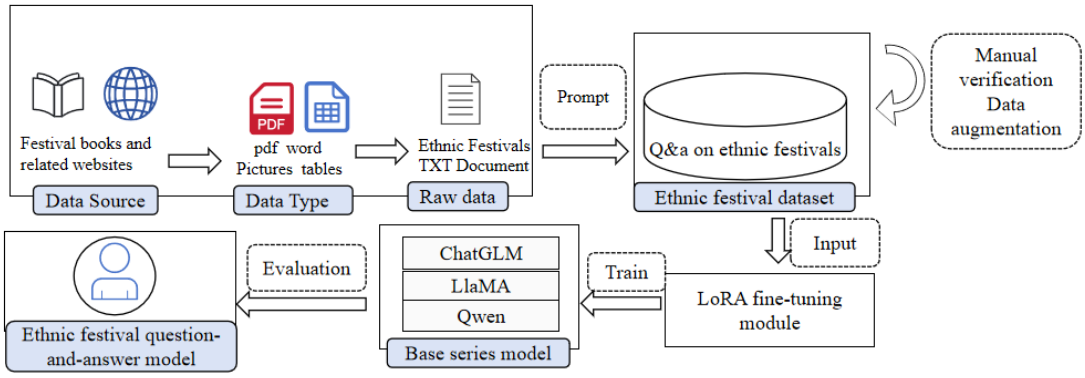

FIGURE 1: OVERALL FRAMEWORK DIAGRAM

2.1 DATASET CONSTRUCTION

China's 56 ethnic groups collectively celebrate over 220 traditional festivals. The Chinese Festival Traditions Reader (Collector's Edition) provides detailed documentation of mainstream Han Chinese festivals, covering foundational information, associated customs, cultural significance, and regional/ethnic characteristics. The Ethnic Minority Festivals compendium meticulously records the festivals of the vast majority of ethnic minorities, organized according to their geographical distribution patterns. During data preprocessing, the content from these books was first OCR-recognized and segmented by "ethnic group" or "festival name," then organized into standardized TXT format documents. To enhance data completeness, the Scrapy crawler framework was used to extract supplementary materials on ethnic festivals from authoritative online platforms like the Ethnic Minorities Network and Baidu Baike. This ultimately yielded a raw corpus of approximately 500,000 characters and over 16,000 question-answer pairs. To enhance the model's understanding of ethnic festival knowledge, key information such as festival names, start dates, customs, and activities within the raw corpus was structurally annotated. Finally, specialized prompt templates were designed based on the annotation results to guide the model in generating relevant questions. The reasoning model was then guided to generate corresponding answers based on the text and annotated content. The dataset construction process is illustrated in Figure 2.

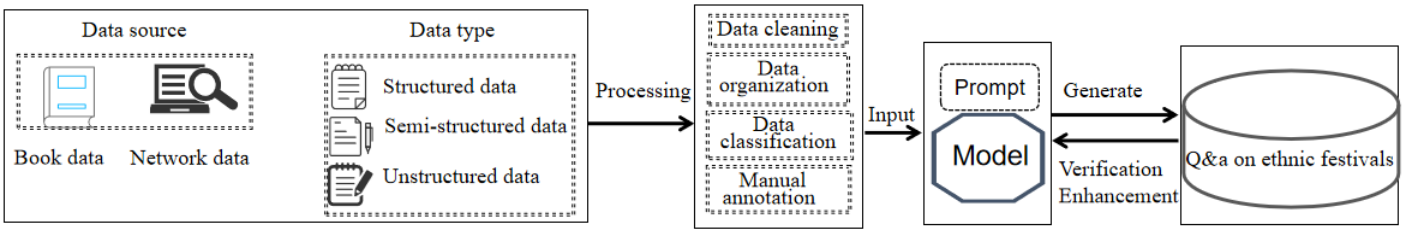

FIGURE 2: FLOWCHART OF DATASET CONSTRUCTION

Data Processing: During the data processing phase, collected data undergoes cleaning first. Duplicate entries are identified and removed using MD5 hash value comparisons to ensure data uniqueness. Manual verification and matching rules filter out non-standard, erroneous, and irrelevant information. Subsequently, festival descriptions undergo semantic normalization (e.g., standardizing "Han ethnic group" and "Han" to consistent terminology), transforming raw data into a standardized format.

Manual Annotation: Manual annotation is a critical step in this research. By annotating key information such as festival names, dates, customs, food traditions, activities, and locations within ethnic festivals, a structured knowledge system for ethnic festivals is constructed. This transforms unstructured text into a machine-parsable format. This work not only enhances the accuracy and reliability of the constructed question-answer pairs but also provides precise supervision signals for subsequent LoRA technology.

Prompt-Based Q&A Dataset Generation: This study employs a question-answer separation generation strategy during the construction of the ethnic festival Q&A dataset. By independently designing prompts for question generation and

Journal of http://www.satursonpublishing.com/
Research in Multidisciplinary Methods and Applications
ISSN 2957-3920 (Online) ISSN 3007-7060 (Print)
Volume 4, Issue 10, October 2025

answer generation, precise control over question types and answer quality is achieved, effectively avoiding cross-contamination between questions and answers.

Data Augmentation:Data augmentation is crucial for enhancing model generalization and robustness. This paper employs prompt-guided strategies to semantically reconstruct original question-answer pairs, thereby expanding dataset scale and enhancing content diversity. First, multi-context semantic rewriting ensures question-answer pairs are restated while preserving original meaning. Subsequently, synonym and near-synonym replacement introduces lexical diversity and mitigates overfitting. These methods broaden training corpus coverage and linguistic richness, providing robust data support for model fine-tuning.The format of the ethnic festival Q&A dataset is shown in Table 1. An example of ethnic festival annotation is shown in Table 2.

TABLE 1: ETHNIC FESTIVAL Q&A DATASET FORMAT

Ethnic Festival Q&A Dataset Format
{"instruction": "When was the name “Spring Festival” officially established?" , "input":"On September 27, 1949, at the First Plenary Session of the Chinese People's Political Consultative Conference, it was resolved to adopt the universally recognized Gregorian calendar era. January 1st of the Gregorian calendar was designated as New Year's Day, commonly known as the solar calendar year. The first day of the first lunar month typically falls around the beginning of spring, hence it was designated as the “Spring Festival,” commonly known as the lunar calendar year." "output":"The name “Spring Festival” was formally established on September 27, 1949. At that time, the First Plenary Session of the Chinese People's Political Consultative Conference adopted the Gregorian calendar system, designating January 1 as New Year's Day. The first day of the first lunar month, being close to the beginning of spring, was officially named “Spring Festival,” commonly known as the Lunar New Year. This marked the unified recognition of the Spring Festival as a national ethnic festival at the state level, a designation that continues to this day."}

TABLE 2: EXAMPLES OF ANNOTATIONS FOR ETHNIC FESTIVALS

Festival	Tim	Custom	Food	activit	Stor
Qingming Festival	4.4	Tomb-sweeping	Wu Ren Fan	Tug-of-war	Wu Ren Fan

2.2 LORA-BASED MODEL FINE-TUNING METHOD

Low-Rank Adaptive Refinement (LoRA) fine-tuning technology embeds trainable low-rank decomposition matrices into each layer of pre-trained models while freezing the original backbone network's weight parameters, significantly reducing the number of parameters requiring training. This method enhances training efficiency, mitigates overfitting risks, and enables domain-specific parameter adjustments to meet specialized vertical task requirements.During fine-tuning, the pre-trained parameter update expression is$W+\Delta W$, with the total parameter update size being $O(d\times d)$, consuming substantial resources.LoRA employs low-rank decomposition approximation for updates, transforming the forward propagation into $h=W_0x+BAx$. This reduces the parameter size to $O(r\times(d+d))$ while freezing the original weight matrix and updating only the A and B parameters, thereby preserving the original structural knowledge. This study employs the LoRA strategy to enhance the large model's capabilities in areas such as ethnic festival culture, thereby improving its performance in vertical domain question-answering tasks.

3 EXPERIMENTS AND RESULTS ANALYSIS

3.1 EVALUATION METRICS

This study employs the BLEU [7]and ROUGE[8] metrics, widely used in natural language processing, as evaluation tools. Both metrics achieve objective quantitative assessment of generated text's surface form by calculating n-gram overlap between generated and reference texts, with BLEU emphasizing accuracy and ROUGE emphasizing recall.

The core concept of BLEU is to measure the consistency between generated and reference content by statistically evaluating the overlap rate of n-grams (sequences of n consecutive words) between them. It also incorporates a sentence-breaking penalty mechanism to prevent the illusion of high accuracy caused by generating excessively short texts. A higher BLEU score indicates greater similarity between the generated text and the reference text. Its calculation formula is given by Equation (1).

$$\mathrm{BLEU-N} = \mathrm{BP}\cdot\exp\left(\sum_{n=1}^{N} w_n \log p_n\right) \quad (1)$$

Among these, N represents the maximum sequence length considered (typically set to 4), while n-gram accuracy denotes the number of n-gram matches between the generated text and the reference text. The weight is assigned to each n-gram, and BP serves as the length penalty factor, calculated according to Equation (2). Here, L_m denotes the length of the model-generated translation, and L_(ref) represents the optimal matching length of the reference translation.

$$BP = \begin{cases} 1 & if \ \ l_b < l_a \\ exp(1-\frac{l_a}{l_b}) & if \ \ l_b \geq l_a \end{cases} \quad (2)$$

Journal of http://www.satursonpublishing.com/
Research in Multidisciplinary Methods and Applications
ISSN 2957-3920 (Online) ISSN 3007-7060 (Print)
Volume 4, Issue 10, October 2025

ROUGE primarily measures the completeness of coverage of content and information in generated text relative to reference text. Its core principle involves assessing text generation quality by calculating n-gram co-occurrence between generated and reference texts. ROUGE-N is computed by splitting both generated and reference outputs into multiple n-grams and applying formula (3). ROUGE-L measures sentence-level semantic similarity by identifying the longest common subsequences between the generated text and the reference text. Its calculation formula is given by Equation (4).

$$\text{ROUGE} - \text{N} = \frac{\sum_{S\in\{\text{ReferemceSummaries}\}}\sum_{\text{gram}_n\in S}\text{Count}_{\text{match}}(\text{gram}_n)}{\sum_{S\in\{\text{ReferemceSummaries}\}}\sum_{\text{gram}_n\in S}\text{Count}(\text{gram}_n)} \quad (3)$$

$$\text{ROUGE} - \text{L} = \frac{\sum_{S\in\{\text{ReferemceSummaries}\}} LCS(Reference, Candidate)}{\sum_{S\in\{\text{ReferemceSummaries}\}} length(Reference)} \quad (4)$$

In formula (3), represents the number of n-grams appearing in both the reference text and the generated text, while C denotes the number of n-grams appearing in the reference text. In formula (4), is the length of the longest common subsequence between the reference text and the generated text, and L is the length of the reference text.

3.2 Experimental Environment

This experiment was conducted on an AutoDL server equipped with a 20 vCPU Intel® Xeon® Platinum 8458P central processing unit, an NVIDIA H800 GPU with 80GB of graphics memory, 100GB of RAM, and 30GB of system storage space. The software environment comprised the Ubuntu 22.04 operating system, the PyTorch 2.3.0 deep learning framework, and Python 3.10.

3.3 Comparative Experiments

Prior to experimentation, to ensure model reliability and suitability, the system systematically reviewed literature on mainstream open-source large language models. Using publicly available benchmark performance as the selection criterion, ChatGLM[9], LlaMA, and Qwen[10] series models were chosen as foundational models. After comprehensive comparison of their fine-tuned performance, the ChatGLM4-9B model combined with the LoRA method was ultimately selected as the most suitable for ethnic festival-related question-answering tasks.The comparison results are shown in Table 3.

Table 3: Comparative Experimental Results

Model Name	BLEU-4	ROUGE-1	ROUGE-2	ROUGE-L
ChatGLM4-9B+LoRA	21.36	42.43	22.69	35.25
Llama3.1-8B+LoRA	7.42	35.33	13.45	25.33
Qwen2.5-7B+LoRA	12.43	39.30	19.40	28.83

3.4 Experimental Results and Analysis

Comparative experiments demonstrate that ChatGLM4 combined with LoRA outperforms Llama3.1 and Qwen2.5 models across all metrics in the ethnic festival domain. Metrics such as BLEU-4 reached 21.36, representing improvements of 13.94 points over Llama3.1-8[11]+LoRA and 8.93 points over Qwen2.5+LoRA. To further validate the model's strengths in the niche field of ethnic festivals, questions were randomly selected from the test set for analysis.Table 5 presents the responses to "instruction" questions from the fine-tuned models. ChatGLM4-9B+LoRA demonstrated optimal performance in ethnic festival Q&A, providing accurate, comprehensive, and concise answers. Qwen2.5-7B exhibited factual errors and redundant content after fine-tuning, resulting in low accuracy. Llama3.1-8B performed worst after fine-tuning, showing content confusion and low coverage, failing to meet the requirements for question-answering tasks.

3.5 Ablation Studies

To further validate the effectiveness of LoRA fine-tuning in model experiments on the Ethnic Festival dataset, this paper designed ablation experiments. The ChatGLM series models were selected as the baseline, with three versions established. The original question-answering dataset was partitioned using stratified sampling. Performance evaluation metrics included BLEU and ROUGE series scores. The experimental results are shown in Table 4.

Table 4: Ablation Experiment Results

Model Name	BLEU-4	ROUGE-1	ROUGE-2	ROUGE-L
ChatGLM4-9B+LoRA	21.36	42.43	22.69	35.25
ChatGLM4-9B	4.62	28.18	9.21	19.25
ChatGLM3-6B	3.91	25.12	8.12	18.45

Table 4 results demonstrate that compared to ChatGLM3-6B and ChatGLM4-9B models, the ChatGLM4-9B model trained with LoRA achieves significant improvements across metrics on the ethnic festival question-answering dataset. Compared to the ChatGLM4-9B model, our method achieves improvements of 16.74 points on metrics like BLEU-4. This indicates that the fine-tuned model demonstrates notable enhancements in answer completeness, coherence, and similarity to reference answers, validating the practicality and accuracy of incorporating LoRA training for models in the ethnic festival domain.

Table 5: Sample Q&A Results

Journal of http://www.satursonpublishing.com/
Research in Multidisciplinary Methods and Applications
ISSN 2957-3920 (Online) ISSN 3007-7060 (Print)
Volume 4, Issue 10, October 2025

“Instruction”	“What are the core activities of the Mongolian White Festival?”
ChatGLM4-9B+LoRA answer	The core activities of the Mongolian White Festival primarily include: \n1. Farewell-to-the-Year Toast: Younger generations offer toasts to elders, expressing respect and blessings. \n2. Whole Lamb Banquet and Nine-Dish Dairy Feast: A grand banquet is held, featuring a whole lamb feast complemented by nine distinct dairy dishes, symbolizing reverence for the color white. \n3. Fire-Busi Singing Vigil: People perform songs accompanied by the Fire-Busi (a traditional Mongolian plucked string instrument), staying awake together to welcome the New Year. \nThese activities collectively form the core rituals of the White Festival, reflecting the Mongolian people's emphasis on family reunion and New Year celebrations.
Qwen2.5-7B+LoRA answer	The Mongolian White Festival (also known as the “Milky White New Year”) features a series of core activities embodying prayers for purity, renewal, and bountiful harvests. Key traditions include: - Held on the first full moon after the Lunar New Year, typically between February 12th and 14th in the Gregorian calendar; - Younger generations apply white cream to elders' faces as a gesture of respect and blessing, symbolizing familial harmony; Families prepare a whole lamb feast and nine dishes of dairy products, representing abundance and purity respectively; white lamps are lit during the festival to pray for brightness and purity, symbolizing hope and the start of the new year; families stay up all night singing and dancing to strengthen bonds and express warmth and joy. The White Festival activities demonstrate the importance placed on purity, serving as a vital means to celebrate the new year and pass down culture, reflecting the Mongolian people's aspirations for a better life.
Llama3.1-8B+LoRA answer	The core activities of the Mongolian Bai Festival mainly include the following three items: 1. A whole lamb feast, with lamb as the staple food, symbolizing reunion and harvest; 2. The ceremony of worshipping Heaven with mare's milk wine, where people pray for favorable weather and a bountiful harvest by offering sacrifices to Heaven. 3. Lasso skills competition, showcasing the traditional Mongolian riding techniques and skills. These activities reflect the cultural connotation of the Bai Festival, emphasizing ethnic traditions and community unity.

4 CONCLUSIONS

Addressing the lack of specialized knowledge and insufficient content rigor in current general-purpose large language models within the ethnic festival domain, this study independently constructed a multi-source ethnic festival knowledge question-answering dataset. Multiple models were fine-tuned based on this dataset. By comparing evaluation metrics across fine-tuned models, the optimal ethnic festival knowledge question-answering model was ultimately identified. Ablation experiments further validated the effectiveness of the LoRA fine-tuning strategy in enhancing model adaptability. Overall, the proposed model significantly improves the rigor and expertise of generated answers related to ethnic festivals, demonstrating potential for practical deployment in digital cultural applications.

Although this study has completed foundational exploration in ethnic festival Q&A, future work can deepen in several directions to advance the long-term goal of digital cultural development: First, continuously expand and mine high-quality folklore materials to enhance the dataset's scale and coverage breadth. Second, explore integrating multimodal technologies—combining images and audio—to deliver more vivid and concrete knowledge services.

REFERENCES

[1]Zhang J, Zhou Z. Digital Preservation and Presentation of China's Outstanding Traditional Culture. Journal of Guizhou Minzu University (Philosophy and Social Sciences Edition), online first,1–16 (2025).

[2]Xu Yuemei, Hu Ling, Zhao Jiayi, et al.Technical Application Prospects and Risk Challenges of Large Language Models [J].Computer Applications,2024,44(06):1655-1662.

[3]Peng Shaoliang, Liu Wenjuan. Computing Power Requirements and Future Challenges for Large Models [J/OL]. Communications of the China Computer Federation, 2024(2).

[4]Huang L, Yu W J, Ma W T,et al.A Survey on Hallucination in Large Language Models:Principles,Taxonomy,Challenges,and Open Questions[EB/OL].(2024-11-19).

[5]Yan Jingqun, Readings on Traditional Chinese Festival Culture (Collector's Edition) [M]. Beijing: Oriental Publishing House, 2009.

[6]Ji Chengqian, Ethnic Minority Festivals [M]. Beijing:China Social Sciences Press,2006.

[7]Papineni K,Roukos S,Ward T,et al.BLEU: a method for automatic evaluation of machine translation[J/OL].Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics,2002:311-318.

[8]Lin C Y.Rouge:A package for automatic evaluation of summaries.Association for Computational Linguistics.2004:74-81.

ISSN 2957-3920 (Online) ISSN 3007-7060 (Print)
Volume 4, Issue 10, October 2025

[9]Team GLM,Zeng A H,Xu B,et al.ChatGLM:A Family of Large Language Models from GLM-130B to GLM-4 All Tools[EB/OL].(2024-07-30)[2025-09-06].

[10]Yang A,Yang B S,Hui B Y,et al.Qwen2 Technical Report[EB/OL].(2024-09-10)[2025-09-06].

[11]WANG S,ZHENG Y,WANG G,et al.Llama3-8B-Chinesechat[EB/OL].(2024-05-09)[2025-09-06]