

ViHerbQA: A Robust QA Model for Vietnamese Traditional Herbal Medicine

Quyen Truong, Long Nguyen, Dien Dinh

Proceedings of the 38th Pacific Asia Conference on
Language, Information and Computation (PACLIC 38)

Nathaniel Oco, Shirley N. Dita, Ariane Macalinga Borlongan, Jong-Bok Kim (eds.)

2024

© 2024. Quyen Truong, Long Nguyen, Dien Dinh. ViHerbQA: A Robust QA Model for Vietnamese Traditional Herbal Medicine. In Nathaniel Oco, Shirley N. Dita, Ariane Macalinga Borlongan, Jong-Bok Kim (eds.), *Proceedings of the 38th Pacific Asia Conference on Language, Information and Computation* (PACLIC 38), 449-466. Institute for the Study of Language and Information, Kyung Hee University. This work is licensed under the Creative Commons Attribution 4.0 International License.

ViHerbQA: A Robust QA Model for Vietnamese Traditional Herbal Medicine

Quyen Truong^{1,2}, Long Nguyen^{1,2*}, Dien Dinh^{1,2}

¹Faculty of Information Technology, University of Science, Ho Chi Minh City, Vietnam

²Vietnam National University, Ho Chi Minh City, Vietnam

quyen.uit@gmail.com, {nhblong, ddien}@fit.hcmus.edu.vn

Abstract

This research introduces ViHerbQA¹, a Vietnamese Traditional Herbal Medicine (VTHM) question-answering model. However, the Vietnamese Traditional Herbal Medicine QA dataset is currently unavailable, so we have built a dataset of 208,203 question-answer pairs generated by Large Language Models (LLMs). To ensure quality, we evaluate these pairs using various evaluation metrics. The ViHerbQA model consists of two stages: pre-training and fine-tuning. We pre-train the ViT5 model on the dataset in the pre-training phase for the Open-Book QA task. Then, in the fine-tuning phase, this pre-trained model is used for the Close-Book QA task. The goal of this approach is to enable the model to have a comprehensive understanding of VTHM. We have conducted extensive evaluations comparing it with current Vietnamese QA systems and state-of-the-art LLMs and found that it outperforms them in terms of performance indicators, such as BERTScore alone achieved more than 80%. In comparison, other models underperformed on the ViHerbQA dataset, demonstrating its potential usefulness in Vietnamese herbal medicine research. We also fine-tuned LLama3.1-8B on our dataset and found that it outperformed the other LLMs evaluated in this study.

1 Introduction

Vietnamese Traditional Herbal Medicine (VTHM) is a long-standing tradition with a rich history of medical knowledge. Even though modern medicine has taken over most parts of the world, traditional medicine remains applicable, which shows how valuable it is in terms of research findings and practical applications in medicine. Nevertheless, the lack or nonexistence of specialized artificial intelligence tools like question-answering systems

limits access to useful herbal information necessary for both research and application.

Large language models (LLMs) and natural language processing (NLP) have advanced greatly in recent times. However, these technologies need to be customized to fit traditional medicine, particularly for less common languages. There are many question-answering (QA) systems for English, but they are still rare in Vietnamese, especially in the traditional medicine field. This highlights the significant potential for improving NLP applications in this crucial field.

Some breakthroughs have been realized in question-answering systems construction, including in open domains and specific domains like medicine. Models such as DrQA (Chen et al., 2017), UnitedQA (Cheng et al., 2021) and UniKQA (Oguz et al., 2022), among others, are categorized under open-domain QA because they provide answers to general questions about different topics while BioMedGPT (Luo et al., 2023) and MEDITRON (Chen et al., 2023) fall under medical domain since they are designed specifically to address issues related to health care, using English language. MedChatZH (Tan et al., 2023) is unique in its own way because it represents traditional Chinese medicine through an LLM fine-tuning method based on a dataset derived around ancient Chinese treatments, whereas the Vietnamese ViHealthQA dataset is used in creating SPBERTQA (Nguyen et al., 2022) in Vietnamese medicine field.

Most of these advancements focus on English sources, with few addressing other languages, particularly Vietnamese. This means that there is a need to develop specialized QA systems for Vietnamese, and in this work, we will develop ViHerbQA, a question-answering model for the VTHM field.

ViHerbQA is a robust question-answering model designed to bridge modern technology and traditional knowledge in Vietnamese herbal medicine.

*Corresponding author.

¹Code availability: <https://github.com/queenley/ViHerbQA>

Our research tackles two main problems, which include the lack of a VTHM QA dataset and the need for an operational model that can understand VTHM knowledge. The aim of ViHerbQA is to overcome these challenges by employing the latest techniques that are applicable within VTHM.

The development process of ViHerbQA involves several steps outlined below:

1. **Dataset Creation:** Since no existing VTHM QA dataset has been discovered, we created a new one consisting of 208,203 question-answer pairs using advanced language models such as GPT-3.5, GPT-4o-mini, Gemini Flash, and Gemini Pro because they have high-performance levels and they also support multiple languages.

2. **Dataset Evaluation:** We used evaluation metrics like Semantic Similarity, BLEU (Papineni et al., 2002), and Rouge (Lin, 2004) to validate the dataset’s applicability.

3. **Model Development:** There are two stages when developing ViHerbQA: a) Pre-training: We use ViT5 (Phan et al., 2022a) to train the Open-Book QA task using our own dataset with the expectation that the model can learn comprehensive VTHM knowledge through contexts in the training process. b) Fine-tuning: This stage entails adapting the pre-trained model towards the Close-Book QA task to enable it to give accurate responses.

4. **Performance Evaluation:** To evaluate ViHerbQA’s performance in answering VTHM-related questions, we compared it against other Vietnamese QA models and state-of-the-art LLMs using BLEU (Papineni et al., 2002), ROUGE (Lin, 2004), and BERT-SCORE (Zhang et al., 2019).

This research creates a more robust VTHM model than previous QA systems. The idea behind this approach is to utilize contemporary natural language processing techniques with an under-researched language to build a stronger base for future question-answering for unique domains. With ViHerbQA, we hope to broaden awareness and understanding surrounding VTHMs among scholars, practitioners, and interested individuals.

2 Related works

Over the past few years, we have seen tremendous advancements in question-answering systems, largely due to progress in natural language processing (NLP) and Large Language Models (LLMs). Different domains have received various contributions, but traditional medicine needs to be explored.

Advancements in large language models (LLMs) and effective pre-training techniques have led to the rise in using QA models. BERT (Devlin et al., 2018), XLNet (Yang et al., 2019), and T5 (Raffel et al., 2020) are examples of such models that were tested on benchmark datasets. These models are trained on general text collections and then fine-tuned on specific QA datasets to boost their performance in different QA tasks. For our work, we use ViT5 (Phan et al., 2022b), which is a Vietnamese-adapted version of T5.

Though NLP has achieved significant milestones in English, Vietnamese remains under-studied. Researchers face challenges such as insufficient annotated resources and complex linguistic features exhibited by this language. Despite this, there have been efforts to develop Vietnamese pre-trained language models, including PhoBERT (Nguyen and Tuan Nguyen, 2020), BartPho (Tran et al., 2021), ViT5 (Phan et al., 2022b), PhoGPT (Nguyen et al., 2023), or multilingual ones like XLM-R (Conneau et al., 2019) and mBERT (Pires et al., 2019) which have shown promising results for less common language NLP tasks. Therefore, this project contributes to Vietnamese NLP through the introduction of a new VTHM QA dataset as well as a robust QA model tailored specifically for this area.

The involvement of AI in healthcare is growing at an unprecedented pace with a focus on drug discovery, disease diagnosis, personal health care, and others (Ching et al., 2018). Nevertheless, traditional medicine is also beginning to find its way into AI systems despite being in nascent stages (Wu et al., 2022); some instances where machine learning has been employed include predicting Chinese medicinal herb components (Han et al., 2018) and identification of possible drug-herbs interactions (Tatonetti et al., 2012). However, more research is needed to create AI-powered question-answering systems for traditional medicine like Vietnamese herbal medicine. This paper fills that void by presenting ViHerbQA, a specialized QA model that aims to provide necessary VTHM knowledge access and comprehension.

By developing a specific question-answering model on VTHM, we can combine Indigenous Traditional Medicine knowledge with state-of-the-art natural language processing techniques. We address the absence of a dedicated VTHM QA dataset by creating one using LLMs with advanced prompting methods and rigorous evaluation procedures. Furthermore, we demonstrate how powerful Viet-

namese language models can be adapted to domains such as VTHM through pre-training and fine-tuning strategies.

3 Dataset

The ViHerbQA dataset is a vast and intricate resource that has been designed to create and evaluate QA systems focused on VTHM. It consists of 208,203 samples composed of various large language models (LLMs) like GPT4o-mini, Gemini-Pro, Gemini-Flash, and GPT-3.5. These samples are composed of question-and-answer pairs that have been generated by LLMs from herbal articles to ensure a broad coverage of VTHM themes. To assess the quality and usability of the dataset, we use BLEU (Papineni et al., 2002), ROUGE (Lin, 2004) as well as semantic similarity at word level metrics.

3.1 Generation

Gemini-Flash, Gemini-Pro, and GPT-3.5 generate the ViHerbQA’s dataset by drawing information from articles about traditional medicinal herbs and formulating question-answer pairs as illustrated in Figure 1.

1. Article Crawling and Pre-processing: We collected 745 herbal articles from YouMed², which is a reliable source that offers comprehensive medical knowledge in Vietnam. These articles undergo a thorough pre-processing procedure such as cleaning of data, normalization, extraction of useful herbal information such as title (herb’s name), description, chemical composition, use, application methods, researches carried out on it, taboo actions and remedies to make sure that just relevant data goes to CSV format. It provides a firm foundation for generating the question-answer pairs.

2. Question Generation: The preprocessed data for each herb’s article falls into two main categories: the entire content and only the herbal prescriptions content. The LLMs, including Gemini-Flash, Gemini-Pro, and GPT-3.5, use the entire content to create 20 questions. This comprehensive approach allows the LLMs to synthesize all the knowledge in the article, resulting in comprehensive questions. For the context of medicinal remedies, an additional five questions are generated by LLMs based on the content of the provided herbal prescriptions. When using LLMs, we ap-

ply unique prompting techniques like Zero-Shot and Few-Shot. We use multiple temperature values to create more diversity from beginner to advanced. Questions are created in this manner and then reviewed to eliminate any redundant characters that may exist before saving the text file format with each question located on 1 line of the file, facilitating the creation of answers in the next step is convenient. The prompt template used for this Question Generation step is provided in Table 1. In this step, we use Few Shot prompt techniques with two prompt versions: one for the entire content and one for the prescription content. In the prompt template, {doc} represents the knowledge content while {herb} denotes the herb’s name. The sum of tokens for this step is approximately 2500 for each query. An example of the output is shown in Figure 3 in Appendix A when the entire content is fed into LLMs, and another example is displayed in Figure 5 in Appendix A when the herbal prescriptions content is used.

3. Answer Generation: We have generated answers corresponding to our prepared questions by feeding information from herbal articles into LLMs. Each question is strongly linked to its corresponding article content, ensuring that the responses are accurate and meaningful. In addition to answering generation, we prompt LLMs to produce the relevant context derived from the original article; the context should be one of the sections or subsections in the article, supporting the Open-Book model training stage. Like in the above stage, a thorough check is performed after generating questions using LLMs. The criteria met to ensure question-answer pairs are included in ViHerbQA’s training dataset include: must be Vietnamese, must not contain meaningless characters, and the answer, question, and context must relate professionally with each other. Therefore, we have employed well-known metrics such as BLEU (Papineni et al., 2002), ROUGE (Lin, 2004), and Semantic-Similarity, among others, for evaluation purposes, which assured us about the quality of the ViHerbQA dataset. Please refer to Figure 6 in Appendix A for the prompt template in this Answer Generation step. Figure 8 depicts the output of this step, with questions created from the entire article content as input. In contrast, Figure 10 shows the output with input consisting of questions related to traditional medicine remedies. We use Vietnamese to design the prompt for this step instead of English because this is essential; answers from LLMs need

²<https://youmed.vn/tin-tuc/y-hoc-co-truyen/duoc-lieu/>

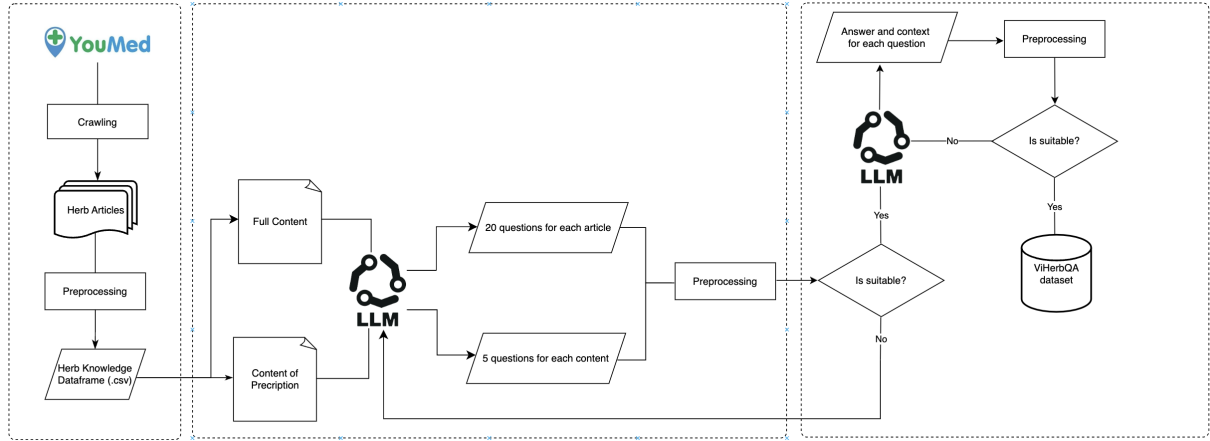


Figure 1: ViHerbQA’s dataset creation pipeline has three stages, including (1) article crawling and pre-processing, (2) question generation, and (3) answer generation.

Prompt template for the entire content	Prompt template the herbal prescriptions content
As a student, please generate 20 Vietnamese questions naturally and generally about <{herb}>, which is a medicinal herb in Vietnamese Traditional Medicine using this knowledge: “{docs}”	As a patient, please generate 5 Vietnamese questions naturally and generally about oriental medicine remedies, specifically, medicinal herbs used in Vietnamese Traditional Medicine. The generated questions should not contain the proper names of any specific herbs, using this knowledge: “{docs}”

Table 1: The prompt template for the Question Generation step.

	Min word count	Max word count
Question	4	95
Answer	1	743
Context	10	2720

Table 2: Statistics on the number of words in questions, answers, and contexts.

to be accurate and highly practical based on the context content provided. Suitable for real situations. For that reason, with the desire for LLMs to be able to read, understand, and extract information from Vietnamese medicinal texts correctly, we use prompts in Vietnamese to have language uniformity to avoid confusion between languages in providing answers to LLMs. The sum tokens for this step are almost 6000 tokens for each query.

We perform statistical analysis steps before we evaluate the dataset. These steps give an overview of the language in the dataset. The data set spans various language levels as illustrated in Table 2 ranging from elementary to most proficient. This diversity is crucial since it makes the dataset resemble real-life settings, thus making post-trained models more realistic.

This ensures that essential traditional medicine

terminology is preserved in the generated dataset. We do this by visualizing the frequency of nouns in the initial articles and within the dataset post-generation. Figure 2 illustrates that the dataset successfully maintains critical terms from the traditional medicine field, such as herb (cây thuốc), medicinal herb (dược liệu), medicinal taste (vị thuốc), traditional medicine (y học cổ truyền), and doctor (bác sĩ).

3.2 Evaluation

Since we generated this dataset using LLMs, there is no gold answer to evaluate. Therefore, we provide another suitable evaluation method based on the word similarity between the generated dataset and the input articles. We can use this method to evaluate various datasets.

Our evaluation process begins with using the Pyvi library³, a robust tool that attaches a PosTag for Vietnamese to each tokenizer found in articles, generated questions and generated answers. We then meticulously focus our evaluation on nouns, verbs, and adjectives. These keyword types demonstrate the accuracy of the generated dataset in containing relevant traditional medical knowledge,

³<https://github.com/trungtv/pyvi>



Figure 2: Visualization of WordCloud including (a) WordCloud of Articles from Youmed, (b) WordCloud of Questions generated by LLMs, and (c) WordCloud of Answers generated by LLMs.

making them the most suitable for our evaluation. We evaluate the generated questions and answers individually and then compute the weighted average of the questions and answers.

We use BLEU, ROUGE, and Semantic Similarity to calculate the similarity between articles and the generated dataset by measuring word-level similarity with nouns (*n*), verbs (*v*), and adjectives (*adj*). Semantic Similarity is calculated by using the FastText library (Joulin et al., 2016) for Vietnamese⁴ to get embedding of each word, then using the cosine similarity measure to calculate similarity.

Let M be the metric used for calculation (M is one of the three mentioned metrics). M_{qa} is the evaluation result of M for the ViHerbQA’s dataset, which is computed as the formula 1, in that, M_n , M_v , and M_{adj} are the similarity measures of nouns, verbs, and adjectives.

$$\frac{2 \times M_n + M_v + M_{adj}}{4} \quad (1)$$

We assign a higher weight to nouns than verbs and adjectives because, in terms of language, the similarity of nouns in this field can prove the expert similarity to specialized words like traditional medicine, doctors, and herbalists. Based on the evaluation method we have provided, the results shown in Table 3 show a high level of similarity in both vocabulary and semantics between articles and ViHerbQA’s dataset when reaching BLEU 60.26, ROUGE 55.82, and Semantic Similarity 99.58.

After thoroughly cleaning, testing, and evaluating, the ViHerbQA dataset has been partitioned into three subsets: training, validation, and testing,

BLEU	ROUGE	Semantic Similarity
60.26	55.82	99.58

Table 3: ViHerbQA’s dataset evaluation on three metrics, including BLEU, ROUGE, and Semantic Similarity.

Training	Validation	Testing
145,742	56,214	6,247

Table 4: The number of question-answer pairs in each training, validation, and testing set of the ViHerbQA dataset.

with a distribution ratio of 7:2:1. Table 4 presents each subset’s corresponding number of question-answer pairs. Table 5 illustrates some samples of the ViHerbQA dataset.

4 Experiments

In our research, we develop a novel two-phase approach to training QA models. The first phase, which we call the “Open Book” phase, takes advantage of the model’s ability to retrieve relevant information from a given context. This simulates the situation of a student consulting reference materials during an exam.

The subsequent stage would be what we refer to as the “Closed Book” phase. At this point, the model is taught to give answers based on its own knowledge obtained from the earlier phase and without any contextual information. This phase mimics the retrieval of previously learned material in the form of a closed-book exam.

This two-stage approach has been particularly successful in answering the specific problems of

⁴<https://huggingface.co/facebook/fasttext-vi-vectors>

Herb	Question	Answer	Context
Kim ngân: Vị thuốc quý với công dụng bất ngờ (Kim ngân: A valuable herb with surprising effects)	Bác sĩ ơi, có phương pháp nào trong y học cổ truyền giúp giảm triệu chứng sốt cảm và mụn nhọt không? (Doctor, is there any traditional medicine method that helps reduce cold symptoms and acne?)	Bài thuốc Ngân kiều tán chứa Kim ngân hoa và Liên kiều có thể được sử dụng để chữa mụn nhọt và sốt cảm. Cách dùng là mỗi vị 40g Kim ngân hoa và Liên kiều, cùng với các vị khác như Kinh giới tuệ, Cát cánh, Đạm đậu sị, Bạc hà, Ngưu bàng tử, và Đạm trúc điệp, tất cả sấy khô tán bột, có thể làm thành viên và uống 1 – 2 lần mỗi ngày, mỗi lần 12g bột. (The “Ngân kiều tán” formula containing Kim ngân hoa (Honeysuckle) and Liên kiều (Forsythia) can be used to treat acne and cold symptoms. The method: use 40g of each Kim ngân hoa and Liên kiều, along with other herbs like Kinh giới tuệ, Cát cánh, Đạm đậu sị, Bạc hà, Ngưu bàng tử, and Đạm trúc điệp, all dried and powdered. It can be made into pills, and take 1-2 times daily, with each dose being 12g of powder.)	Bài thuốc Ngân kiều tán Thường dùng chữa mụn nhọt, sốt cảm: Kim ngân hoa và Liên kiều mỗi vị 40g; Kinh giới tuệ 16g; Cát cánh 24g; Đạm đậu sị 20g; Bạc hà và Ngưu bàng tử mỗi vị 24g; Đạm trúc điệp 15g. Tất cả sấy khô tán bột. Có thể làm thành viên. Uống 1 – 2 lần mỗi ngày, mỗi lần 12g bột.... (The “Ngân kiều tán” formula is commonly used to treat acne and colds: Kim ngân hoa and Liên kiều, 40g each; Kinh giới tuệ 16g; Cát cánh 24g; Đạm đậu sị 20g; Bạc hà and Ngưu bàng tử, 24g each; Đạm trúc điệp 15g. All dried and powdered. Can be made into pills. Take 1-2 times daily, 12g per dose...)
Mộc thông: Vị thuốc có công dụng lợi tiểu (Mộc thông: An herb with diuretic properties.)	Trong trường hợp nào cần phải kiêng kị khi sử dụng Mộc thông và những tác động tiêu cực có thể xảy ra? (In what cases should Mộc thông be avoided, and what negative effects can occur?)	Người suy nhược, hoạt tính, mệt mỏi, không có thấp nhiệt bên trong thì cấm dùng. Phụ nữ có thai và những người tiểu tiện quá nhiều không được dùng. (Those who are weak, experiencing spermatorrhea, fatigue, or do not have internal damp-heat should avoid using it. Pregnant women and those who urinate excessively should not use it.)	Người suy nhược, hoạt tính, mệt mỏi, không có thấp nhiệt bên trong thì cấm dùng. Phụ nữ có thai và những người tiểu tiện quá nhiều không được dùng. Mộc thông có công dụng lợi tiểu tiện, trị bí tiểu, tiểu gắt buốt, ngoài ra còn bị bế kinh. Những thông tin trên đây chỉ mang tính chất tham khảo. Tốt nhất, bạn cần tham khảo ý kiến bác sĩ trước khi sử dụng các vị thuốc!... (Those who are weak, experiencing spermatorrhea, fatigue, or do not have internal damp-heat should avoid using it. Pregnant women and those who urinate excessively should not use it. Mộc thông has diuretic properties, treating urinary retention and painful urination, and can also address amenorrhea. The information provided here is for reference only. It is best to consult a doctor before using any herbal medicines!...)
Vông nem: Thứ lá cây chữa mất ngủ hiệu quả (Vông nem: A leaf that effectively treats insomnia)	Vông nem có tác dụng an thần như thế nào? (How does Vông nem have a sedative effect?)	Vông nem có tác dụng an thần nhờ vào các thành phần hóa học có trong lá, giúp làm dịu hệ thần kinh. (Vông nem has a sedative effect thanks to the chemical components in its leaves, which help to calm the nervous system.)	Lá Vông nem có tác dụng ức chế hệ thần kinh trung ương, làm yên tĩnh, gây ngủ, hạ nhiệt, hạ huyết áp...(Vông nem leaves have effects that inhibit the central nervous system, inducing calmness, sleep, reducing fever, and lowering blood pressure...)

Table 5: Some samples of ViHerbQA dataset.

VTHM. With this approach, our model allows high and complex comprehension of VTHM while being able to give exact and trusted feedback without any context.

4.1 Model

The development of ViHerbQA begins with pre-training the ViT5 model (Phan et al., 2022a) for the Open-Book task using the constructed dataset. Preparing the model to answer questions within the context of the Open-Book task will enable exhaustive learning of VTHM’s knowledge from the provided contexts. Given that VTHM necessitates high exactness and trustworthiness, the Open-Book

pre-trained model will enhance the capability to supply more accurate answers for the Close-Book task in the subsequent step.

Upon reaching convergence in training with the Open-Book task, we will employ that model to fine-tune for the Close-Book task. We leverage a VTHM knowledge model for fine-tuning, assuring ViHerbQA delivers more exact, experienced answers without needing the context for this field.

We use ViT5 (Phan et al., 2022a) to take advantage of the architecture’s encoding and decoding capabilities and the model’s ability to understand Vietnamese to serve as a suitable basis for training ViHerbQA, a QA model serving the Vietnamese

	BLEU	Rouge1	Rouge2	RougeL	RougeLsum	BertScore (P)	BertScore (R)	BertScore (F1)
Large Language Models								
<i>GPT_{3.5}</i>	6.19	31.28	17.62	23.40	24.99	66.13	74.34	69.93
<i>Gemini_{flash}</i>	3.42	22.41	12.79	17.19	19.14	60.37	71.92	65.56
<i>Llama3.1-8B</i>	6.24	31.32	15.59	22.51	22.81	64.27	71.42	67.48
Vietnamese Transformer Models								
<i>ViT5_{base}</i>	4.77	24.17	11.35	19.11	19.15	60.54	60.81	60.58
<i>ViT5_{large}</i>	0.51	2.43	0.48	2.28	2.28	38.91	32.99	35.66
<i>BartPho_{base}</i>	14.26	49.35	29.64	38.30	38.37	77.03	70.28	73.43
<i>BartPho</i>	14.15	49.26	29.46	38.19	38.27	77.00	70.28	73.42
ViHerbQA (our)								
<i>OpenBook_{base}</i>	31.72	59.90	47.74	52.84	52.84	85.57	77.78	81.36
<i>OpenBook_{large}</i>	33.00	60.47	49.08	53.83	53.85	86.14	78.24	81.87
<i>CloseBook_{base}</i>	31.17	59.44	46.98	52.19	52.21	85.22	77.60	81.11
<i>CloseBook_{large}</i>	32.43	60.01	48.30	53.20	53.21	85.78	78.06	81.62
<i>Llama3.1-8B_{ft}</i>	8.80	37.54	20.41	26.70	27.53	65.11	74.97	69.38

Table 6: Evaluation of ViHerbQA compared to other models based on BLEU, Rouge, and BertScore metrics.

Traditional Herb Medicine field.

We employ RTX 3090 - 24GB VRAM to train ViHerbQA on ViT5-base and ViT5-large instances. The models undergo five epochs of training with a batch size of four for ViT5-base and two for ViT5-large during both the Open-Book and Close-Book phases. We use the Adam optimizer with a learning rate set to 1e-5. For ViT5-base, the Question Max Length and Answer Max Length are 512 and 1024, respectively, while for ViT5-large, they are 256 and 512.

4.2 Fine-tuning Llama3.1-8B

We have fine-tuned the Llama3.1-8B, one of the recent modern LLM models, through two steps: pre-training and fine-tuning. During the first stage, the model is pre-trained on the entire text from the articles about VTHM. This step is essential for the model to grasp the knowledge of VTHM before moving on to the second stage, which is fine-tuning the model for the QA task. The LoRA technique is used for training in both stages. The epochs for the first stage are 15, and the second stage is 2. According to Table 6, after fine-tuning the VTHM dataset, Llama3.1-8B outperforms the original Llama3.1-8B and the Llama3.1 versions with more extensive parameters. However, compared to the model built based on the ViT5 model, the results of Llama3.1-8B still need improvement. This suggests that using a specialized model for Vietnamese would be more effective than a multilanguage LLM.

4.3 Result

We use the test set of the ViHerbQA dataset to evaluate the ViHerbQA model in both the open-

book and close-book stages. We also evaluated the model, which is fine-tuned Llama3.1-8B on our dataset, and compared it with the original Llama3.1-8B. Additionally, we assess two LLMs, GPT3.5 and Gemini Flash, to address concerns about the capability of today’s LLMs to provide accurate answers within almost all domains. Furthermore, we exhaustively consider the ViHerbQA model compared to strong Vietnamese transformer models such as ViT5 (Phan et al., 2022a) and BartPho (Tran et al., 2021).

The results presented in Table 6 demonstrate the exceptional linguistic and semantic capabilities of the ViHerbQA model in the VTHM domain, as evidenced by metrics such as BLEU, ROUGE (including Rouge1, Rouge2, RougeL, and RougeLsum), and BertScore (including BertScore Precision (P), BertScore Recall (R), and BertScore F1) for Vietnamese, with BertScore values greater than 80% for both the base and large versions, the ViHerbQA model shows very high semantic accuracy. Moreover, BLEU values exceeding 30% and ROUGE values surpassing 50% underscore ViHerbQA’s diverse linguistic abilities. Furthermore, when fine-tuning the Llama3.1-8B in our dataset, the results are significantly better than the original LLM, even outperforming other LLMs in almost all metrics.

Refer to Table 7 for some responses from the ViHerbQA model from the ViHerbQA test dataset. The Question column in this table presents the input question fed into the model; the Gold Answer is the answer of the test dataset, and the Predicted Answer is the reply of the ViHerbQA model.

ViT5 and BartPho are two transformer models that were powerfully developed for Vietnamese

Question	Gold Answer	Predicted Answer
Thời gian đắp thuốc từ lá bàng tươi lên vùng da bị bệnh là bao lâu? (<i>How long should the poultice made from fresh leaves be applied to the affected area?</i>)	Thời gian đắp thuốc từ lá bàng tươi lên vùng da bị bệnh là 15 phút. (<i>The poultice made from fresh leaves should be applied to the affected area for 15 minutes.</i>)	Thời gian đắp thuốc từ lá bàng tươi lên vùng da bị bệnh là 15 phút. (<i>The poultice made from fresh leaves should be applied to the affected area for 15 minutes.</i>)
Mộc thông có công dụng chính nào? (<i>What are the main uses of Mộc thông?</i>)	Mộc thông có tác dụng lợi tiểu tiện, thông huyết mạch. (<i>Mộc thông has diuretic properties, treating urinary retention, and regulating blood circulation.</i>)	Mộc thông có tác dụng lợi tiểu tiện, thông huyết mạch. (<i>Mộc thông has diuretic properties, treating urinary retention, and regulating blood circulation.</i>)
Cách sử dụng Bắc độn để chữa tim hồi hộp và miệng khô như thế nào? (<i>How is Bắc độn used to treat shortness of breath and dry mouth?</i>)	Bắc 4 g, Lá tre, Mạch môn mỗi vị 12 g, sắc uống. (<i>Bắc 4g, Bamboo leaves, and Mạch môn 12g each, to be decocted into a drink.</i>)	Bắc 4 g, Lá tre, Mạch môn mỗi vị 12 g, sắc uống. (<i>Bắc 4g, Bamboo leaves, and Mạch môn 12g each, to be decocted into a drink.</i>)
Tinh dầu phong lữ có đặc tính gì có lợi cho da nhờn? (<i>Does Phong lữ (Geranium) oil have any benefits for oily skin?</i>)	Với đặc tính kiểm dầu, tinh dầu phong lữ có thể giúp cân bằng lượng dầu trên da. (<i>With its astringent properties, Phong lữ oil can help balance the skin's oil levels.</i>)	Với đặc tính kiểm dầu, tinh dầu phong lữ có thể giúp cân bằng lượng dầu trên da. (<i>With its astringent properties, Phong lữ oil can help balance the skin's oil levels.</i>)
Các tác dụng phụ có thể gặp khi sử dụng thận xạ là gì? (<i>What side effects might occur when using kidney radiation?</i>)	Khi dùng có thể gây ra các tác dụng phụ như đau bụng, tiêu chảy, đau đầu (<i>It can cause side effects such as abdominal pain, diarrhea, and headaches.</i>)	Khi dùng có thể gây ra các tác dụng phụ như đau bụng, tiêu chảy, đau đầu (<i>It can cause side effects such as abdominal pain, diarrhea, and headaches.</i>)
Liên tu có thể ảnh hưởng đến cơ thể như thế nào trong trường hợp sử dụng dài hạn, đặc biệt là đối với người suy nhược hoặc có tiểu tiện bí? (<i>How can Lotus affect the body in cases of long-term use, especially for those who are debilitated or have urinary retention?</i>)	Cơ thể suy nhược, táo bón, tiểu tiện bí không nên dùng Liên tu. (<i>People with a debilitated body, constipation, or urinary retention should not use Lotus seeds.</i>)	Cơ thể suy nhược, táo bón, tiểu tiện bí không nên dùng Liên tu. (<i>People with a debilitated body, constipation, or urinary retention should not use Lotus seeds.</i>)

Table 7: Some responses of the ViHerbQA model.

only. However, with the result in Table 6, we can see that two models have yet to have the ability to solve questions in VTHM.

The outcomes of GPT3.5 and Gemini Flash show that while state-of-the-art LLMs excel in numerous domains, they have significant limitations in expert domains, particularly VTHM. This unlocks the opportunity for researchers to investigate new approaches that mix modern technologies with traditional wisdom, as demonstrated by ViHerbQA.

5 Conclusion

This work presents ViHerbQA, the first question-answering system for VTHM. The absence of VTHM-specific QA datasets is addressed by using various state-of-the-art LLMs to build a dataset with 208,203 question-answer pairs that are then carefully evaluated using numerous appropriate language evaluation metrics. With its foundation on sturdy ViT5 architecture, we have constructed the ViHerbQA model in two stages, and it performs better than any other model in terms of answering questions about VTHM. Such findings indicate that ViHerbQA outperforms competitive baselines

consisting of state-of-the-art LLMs and existing Vietnamese transformer models, which further emphasize the importance of domain knowledge coupled with fine-tuning for more precise applications within this field. This research contributes valuable resources for the VTHM community and motivates studies combining traditional medicine and modern artificial intelligence.

6 Future work

We will use explainable AI techniques for the ViHerbQA model in the future. These include making visible attention and determining which features are essential to interpreting the model’s responses. Consequently, trust levels in this system may increase, and more understanding about information in VTHM may be offered. In this way, we hope to see big strides in AI-assisted traditional medicine for a more informed application of VTHM into contemporary healthcare systems.

We will also integrate the DPO technique for further enhancement of ViHerbQA in line with users’ preferences. It enables us to adjust the system so that it can produce responses that are more rele-

vant, informative, and culturally sensitive within the domain of VTHM through direct optimization of the model parameters using human feedback.

References

- Danqi Chen, Adam Fisch, Jason Weston, and Antoine Bordes. 2017. Reading wikipedia to answer open-domain questions. *arXiv preprint arXiv:1704.00051*.
- Zeming Chen, Alejandro Hernández Cano, Angelika Romanou, Antoine Bonnet, Kyle Matoba, Francesco Salvi, Matteo Pagliardini, Simin Fan, Andreas Köpf, Amirkeivan Mohtashami, Alexandre Sallinen, Alireza Sakhaeirad, Vinitra Swamy, Igor Krawczuk, Deniz Bayazit, Axel Marmet, Syrielle Montariol, Mary-Anne Hartley, Martin Jaggi, and Antoine Bosselut. 2023. [Meditron-70b: Scaling medical pretraining for large language models](#). *Preprint*, arXiv:2311.16079.
- Hao Cheng, Yelong Shen, Xiaodong Liu, Pengcheng He, Weizhu Chen, and Jianfeng Gao. 2021. Unitedqa: A hybrid approach for open domain question answering. *arXiv preprint arXiv:2101.00178*.
- Travers Ching, Daniel S Himmelstein, Brett K Beaulieu-Jones, Alexandr A Kalinin, Brian T Do, Gregory P Way, Enrico Ferrero, Paul-Michael Agapow, Michael Zietz, Michael M Hoffman, et al. 2018. Opportunities and obstacles for deep learning in biology and medicine. *Journal of the royal society interface*, 15(141):20170387.
- Alexis Conneau, Kartikay Khandelwal, Naman Goyal, Vishrav Chaudhary, Guillaume Wenzek, Francisco Guzmán, Edouard Grave, Myle Ott, Luke Zettlemoyer, and Veselin Stoyanov. 2019. [Unsupervised cross-lingual representation learning at scale](#). *CoRR*, abs/1911.02116.
- Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.
- Ke Han, Lei Zhang, Miao Wang, Rui Zhang, Chunyu Wang, and Chengzhi Zhang. 2018. Prediction methods of herbal compounds in chinese medicinal herbs. *Molecules*, 23(9):2303.
- Armand Joulin, Edouard Grave, Piotr Bojanowski, Matthijs Douze, H  rve J  gou, and Tomas Mikolov. 2016. [Fasttext.zip: Compressing text classification models](#). *Preprint*, arXiv:1612.03651.
- Chin-Yew Lin. 2004. Rouge: A package for automatic evaluation of summaries. In *Text summarization branches out*, pages 74–81.
- Yizhen Luo, Jiahuan Zhang, Siqi Fan, Kai Yang, Yushuai Wu, Mu Qiao, and Zaiqing Nie. 2023. [Biomedgpt: Open multimodal generative pre-trained transformer for biomedicine](#). *Preprint*, arXiv:2308.09442.
- Dat Quoc Nguyen, Linh The Nguyen, Chi Tran, Dung Ngoc Nguyen, Dinh Phung, and Hung Bui. 2023. [PhoGPT: Generative Pre-training for Vietnamese](#). *arXiv preprint*, arXiv:2311.02945.
- Dat Quoc Nguyen and Anh Tuan Nguyen. 2020. [PhoBERT: Pre-trained language models for Vietnamese](#). In *Findings of the Association for Computational Linguistics: EMNLP 2020*, pages 1037–1042, Online. Association for Computational Linguistics.
- Nhung Thi-Hong Nguyen, Phuong Phan-Dieu Ha, Luan Thanh Nguyen, Kiet Van Nguyen, and Ngan Luu-Thuy Nguyen. 2022. [Spbertqa: A two-stage question answering system based on sentence transformers for medical texts](#). *Preprint*, arXiv:2206.09600.
- Barlas Oguz, Xilun Chen, Vladimir Karpukhin, Stan Peshterliev, Dmytro Okhonko, Michael Schlichtkrull, Sonal Gupta, Yashar Mehdad, and Scott Yih. 2022. [UniK-QA: Unified representations of structured and unstructured knowledge for open-domain question answering](#). In *Findings of the Association for Computational Linguistics: NAACL 2022*, pages 1535–1546, Seattle, United States. Association for Computational Linguistics.
- Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu. 2002. Bleu: a method for automatic evaluation of machine translation. In *Proceedings of the 40th annual meeting of the Association for Computational Linguistics*, pages 311–318.
- Long Phan, Hieu Tran, Hieu Nguyen, and Trieu H Trinh. 2022a. [Vit5: Pretrained text-to-text transformer for vietnamese language generation](#). *arXiv preprint arXiv:2205.06457*.
- Long Phan, Hieu Tran, Hieu Nguyen, and Trieu H Trinh. 2022b. [Vit5: Pretrained text-to-text transformer for vietnamese language generation](#). *Preprint*, arXiv:2205.06457.
- Telmo Pires, Eva Schlinger, and Dan Garrette. 2019. [How multilingual is multilingual bert?](#) *CoRR*, abs/1906.01502.
- Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J Liu. 2020. Exploring the limits of transfer learning with a unified text-to-text transformer. *Journal of machine learning research*, 21(140):1–67.
- Yang Tan, Mingchen Li, Zijie Huang, Huiqun Yu, and Guisheng Fan. 2023. [Medchatzh: a better medical adviser learns from better instructions](#). *Preprint*, arXiv:2309.01114.
- Nicholas P Tatonetti, Guy Haskin Fernald, and Russ B Altman. 2012. A novel signal detection algorithm for identifying hidden drug-drug interactions in adverse event reports. *Journal of the American Medical Informatics Association*, 19(1):79–85.

Nguyen Luong Tran, Duong Minh Le, and Dat Quoc Nguyen. 2021. [Bartpho: Pre-trained sequence-to-sequence models for vietnamese](#). *CoRR*, abs/2109.09701.

Chaoyong Wu, Jianxin Chen, Elaine Lai-Han Leung, Hang Chang, and Xu Wang. 2022. Artificial intelligence in traditional medicine.

Zhilin Yang, Zihang Dai, Yiming Yang, Jaime Carbonell, Russ R Salakhutdinov, and Quoc V Le. 2019. Xlnet: Generalized autoregressive pretraining for language understanding. *Advances in neural information processing systems*, 32.

Tianyi Zhang, Varsha Kishore, Felix Wu, Kilian Q Weinberger, and Yoav Artzi. 2019. Bertscore: Evaluating text generation with bert. *arXiv preprint arXiv:1904.09675*.

A Prompt and Output of the Dataset Generation step

1. Actiso làm thế nào giúp bảo vệ gan và giải độc gan? Cơ chế hoạt động của nó là gì?
2. Thành phần hóa học chính của Actiso bao gồm những gì và làm thế nào các thành phần này ảnh hưởng đến sức khỏe của gan và cơ thể?
3. Tác dụng lợi tiểu và lợi mật của Actiso được thể hiện như thế nào trong việc điều trị bệnh phù, thấp khớp và các vấn đề liên quan đến gan?
4. Cách sử dụng và liều dùng Actiso như thế nào để đạt hiệu quả tốt nhất trong điều trị và bảo vệ sức khỏe?
5. Có những nghiên cứu nào đã chứng minh về tác dụng của Actiso trong việc bảo vệ gan, hỗ trợ tiêu hóa, và ngăn ngừa các bệnh lý khác không?
6. Liệu Actiso có thể được sử dụng trong việc hỗ trợ điều trị ung thư và giảm nguy cơ mắc bệnh không? Cơ chế hoạt động của nó là gì?
7. Trong y học cổ truyền, có những lời khuyên cụ thể nào về cách sử dụng và kết hợp Actiso với các loại thực phẩm khác để tối ưu hóa tác dụng của nó?
8. Có những tác dụng phụ nào có thể xảy ra khi sử dụng Actiso trong điều trị dài hạn? Làm thế nào để đối phó với những tác dụng này?
9. Actiso có ảnh hưởng đến sức khỏe xương và tim mạch như thế nào? Có cách nào để tăng cường hiệu quả của việc sử dụng Actiso trong việc bảo vệ hệ xương và tim mạch không?
10. Trong trường hợp cần thiết, liệu có những biện pháp thay thế nào có thể sử dụng để thay thế hoặc bổ sung cho việc sử dụng Actiso?
11. Có những biện pháp nào cần được thực hiện để đảm bảo an toàn và hiệu quả khi sử dụng Actiso, đặc biệt là khi kết hợp với các loại thuốc khác?
12. Có những nghiên cứu nào về tương tác giữa Actiso và các loại thuốc hoặc thực phẩm khác mà chúng ta nên biết?
13. Actiso có thể ảnh hưởng đến quá trình tiêu hóa hay hấp thụ chất dinh dưỡng không? Nếu có, thì làm thế nào?
14. Trong trường hợp sử dụng Actiso để giảm cân, liệu có những biện pháp điều chỉnh chế độ ăn uống và lối sống cần thiết để đạt hiệu quả tốt nhất không?
15. Có nên sử dụng Actiso trong trường hợp tự điều trị các vấn đề sức khỏe nhẹ mà không tham khảo ý kiến của chuyên gia y tế không?
16. Làm thế nào để đảm bảo tuân thủ đúng liều lượng và cách sử dụng khi sử dụng Actiso?
17. Actiso có thể gây ra tình trạng phụ thuộc hay nghiện nếu sử dụng kéo dài không?
18. Có những hạn chế hay rủi ro gì khi sử dụng Actiso, đặc biệt là trong nhóm người dễ bị ảnh hưởng như trẻ em, phụ nữ mang thai hoặc cho con bú, người già, và người mắc các bệnh mãn tính?
19. Liệu có những tác động phụ nào có thể xuất hiện khi sử dụng Actiso trong điều trị dài hạn?
20. Trong y học cổ truyền, liệu có những lời khuyên nào về cách lựa chọn, thu hái và bảo quản Actiso để đảm bảo hiệu quả và an toàn?

Figure 3: Sample output for the Question Generation step with the Entire Content input.

1. *How does Artichoke help protect the liver and detoxify it? What is its mechanism of action?*
2. *What are the main chemical components of Artichoke, and how do these components affect liver health and overall health?*
3. *How are the diuretic and choleretic effects of Artichoke manifested in the treatment of edema, rheumatism, and liver-related issues?*
4. *How should Artichoke be used, and what is the recommended dosage to achieve the best results in treatment and health protection?*
5. *Are there any studies that have proven the effects of Artichoke in protecting the liver, aiding digestion, and preventing other diseases?*
6. *Can Artichoke be used to support cancer treatment and reduce the risk of disease? What is its mechanism of action?*
7. *In traditional medicine, are there specific recommendations on how to use and combine Artichoke with other foods to optimize its effects?*
8. *What side effects might occur when using Artichoke for long-term treatment? How can these side effects be managed?*
9. *How does Artichoke affect bone and cardiovascular health? Is there a way to enhance the effectiveness of using Artichoke in protecting the skeletal and cardiovascular systems?*
10. *In necessary cases, are there any alternative measures that can be used to replace or supplement the use of Artichoke?*
11. *What measures should be taken to ensure safety and effectiveness when using Artichoke, especially when combined with other medications?*
12. *Are there any studies on the interactions between Artichoke and other drugs or foods that we should be aware of?*
13. *Can Artichoke affect the digestion or absorption of nutrients? If so, how?*
14. *In the case of using Artichoke for weight loss, are there necessary dietary and lifestyle adjustments to achieve the best results?*
15. *Is it advisable to use Artichoke for self-treatment of minor health issues without consulting a healthcare professional?*
16. *How can one ensure proper adherence to dosage and usage instructions when using Artichoke?*
17. *Can prolonged use of Artichoke lead to dependence or addiction?*
18. *What limitations or risks are associated with using Artichoke, especially in vulnerable groups such as children, pregnant or breastfeeding women, the elderly, and those with chronic conditions?*
19. *Are there any potential side effects that may appear with long-term use of Artichoke in treatment?*
20. *In traditional medicine, are there any recommendations on how to select, harvest, and preserve Artichoke to ensure effectiveness and safety?*

Figure 4: Sample output for the Question Generation step with the Entire Content input (translated to English).

1. Tôi muốn hỏi về phương pháp truyền thống dùng dược liệu để điều trị bệnh sổ mũi và say nắng, bạn có thể chia sẻ thông tin về điều này không? (*I would like to ask about traditional methods using herbal remedies to treat a runny nose and heatstroke. Can you share information about this?*)
2. Có bài thuốc dân gian nào được sử dụng trong y học cổ truyền để chữa bệnh ho, sốt và sưng amidan không? (*Are there any folk remedies used in traditional medicine to treat coughs, fever, and swollen tonsils?*)
3. Xin bác sĩ tư vấn về việc sử dụng dược liệu từ cây Bướm bạc để điều trị viêm thận, phù và giúp lợi tiểu. (*Please advise on the use of herbal remedies from the Butterfly Pea plant to treat nephritis, edema, and aid diuresis.*)
4. Tôi đang quan tâm đến bài thuốc truyền thống dành cho bệnh khí hư bạch đới, bạn có thể chia sẻ thông tin chi tiết về cách sử dụng không? (*I am interested in traditional remedies for treating leukorrhea (abnormal vaginal discharge). Could you provide detailed information on how to use them?*)
5. Bác sĩ có thể giới thiệu về cách sử dụng dược liệu từ lá cây Mướp và lá cây Bướm bạc để chữa lở loét da không? (*Could you introduce how to use herbal remedies from Luffa leaves and Butterfly Pea leaves to treat skin ulcers?*)

Figure 5: Sample output for the Question Generation step with the Herbal Prescriptions Content input.

Bạn hãy đóng vai trò là một chuyên gia y học cổ truyền Việt Nam, am hiểu sâu rộng về các loại dược liệu. Bạn được cung cấp một văn bản khoa học về dược liệu {duoc_lieu} với danh mục chi tiết. Nhiệm vụ của bạn là đọc kỹ văn bản và trả lời các câu hỏi dựa trên thông tin được cung cấp.

Đầu vào:

- **DANH MỤC:** Danh mục của văn bản khoa học, mỗi mục trên một dòng.
- **CONTEXT:** Nội dung chính của văn bản khoa học về dược liệu.
- **Câu hỏi:** Danh sách các câu hỏi liên quan đến dược liệu, mỗi câu hỏi trên một dòng.

Yêu cầu:

- Trả lời toàn bộ các câu hỏi chỉ dựa trên thông tin từ context.
- Diễn giải câu trả lời từ góc nhìn của chuyên gia y học cổ truyền khi phù hợp.
- Sử dụng thuật ngữ y học cổ truyền Việt Nam khi cần thiết.
- Trích dẫn chính xác phần context liên quan đến câu trả lời.
- Nếu không có đủ thông tin để trả lời, hãy ghi: “Để có được thông tin chính xác, vui lòng liên hệ với bác sĩ chuyên môn.”
- Định dạng đầu ra: JSON, bao gồm một danh sách các từ điển. Mỗi từ điển gồm ba phần tử:
 - question: Câu hỏi đầu vào đã được chuẩn hóa (loại bỏ ký tự thừa, xuống dòng).
 - answer: Câu trả lời đầy đủ và chính xác.
 - knowledge: Tiêu đề của mục trong {danh_muc} được sử dụng để trả lời câu hỏi.

Ví dụ:

```
[
  {
    "question": "Thành phần hóa học của Actiso chứa những hợp chất nào quan trọng?",
    "answer": "Actiso chứa các hợp chất quan trọng như axit caffeic, flavonoid, lacton sesquiterpene,
    ↳ anthocyan",
    "knowledge": "2. Thành phần hóa học chứa trong Actiso"
  },
  {
    "question": "Actiso có tác dụng gì đối với gan?",
    "answer": "Actiso giúp tăng cường chức năng gan như làm tăng bài tiết dịch mật và giảm nồng độ các
    ↳ độc tố",
    "knowledge": "3. Công dụng của Actiso"
  }
]
```

DANH MỤC

{danh_muc}

CONTEXT:

{context}

Câu hỏi

{list_of_questions}

Hãy trả lời các câu hỏi trên theo định dạng JSON đã yêu cầu.

Figure 6: Prompt template for the Answer Generation step.

You will play the role of a traditional Vietnamese medicine expert with deep knowledge of various medicinal herbs. You are provided with a scientific document about the medicinal herb {herb} with a detailed list. Your task is to carefully read the document and answer questions based on the provided information.

Input:

- ***LIST:*** The list of the scientific document, each item on a new line.
- ***CONTEXT:*** The main content of the scientific document about the medicinal herb.
- ***Questions:*** A list of questions related to the medicinal herb, each question on a new line.

Requirements:

- Answer all questions based only on the information from the context.
- Explain the answers from the perspective of a traditional medicine expert when appropriate.
- Use Vietnamese traditional medicine terminology when necessary.
- Accurately cite the relevant part of the context in the answer.
- If there is insufficient information to answer, write: "For accurate information, please consult a specialist."
- Output format: JSON, including a list of dictionaries. Each dictionary contains three elements:
 - question: The standardized input question (removing extraneous characters, line breaks).
 - answer: The full and accurate answer.
 - knowledge: The title of the section in {list} used to answer the question.

Example:

```
[
  {
    "question": "What important compounds are contained in the chemical composition of Artichoke?",
    "answer": "Artichoke contains important compounds such as caffeic acid, flavonoids, sesquiterpene
    ↪ lactones, anthocyanins",
    "knowledge": "2. The chemical composition contained in Artichoke."
  },
  {
    "question": "What are the effects of Artichoke on the liver?",
    "answer": "Artichoke helps enhance liver function, such as increasing bile secretion and reducing toxin
    ↪ levels",
    "knowledge": "3. Uses of Artichoke"
  }
]
```

MENU

{menu}

CONTEXT:

{context}

Questions

{list_of_questions}

Please answer the questions in the required JSON format.

Figure 7: Prompt template for the Answer Generation step (translated to English).

```

[
{
  "question": "Actiso làm thế nào giúp bảo vệ gan và giải độc gan? Cơ chế hoạt động của nó là gì?",
  "answer": "Actiso chứa cynarin và axit caffeoylquinic, là những chất chống oxy hóa có thể bảo
    ↳ vệ gan. Chúng giúp làm giảm nồng độ các độc tố có hại cho gan và có thể giúp sửa chữa
    ↳ các tế bào gan bị tổn thương.",
  "knowledge": "5. Các nghiên cứu về tác dụng của Actiso"
},
{
  "question": "Thành phần hóa học chính của Actiso bao gồm những gì và làm thế nào các thành
    ↳ phần này ảnh hưởng đến sức khỏe của gan và cơ thể?",
  "answer": "Actiso chứa các hợp chất như dẫn xuất axit caffeic, flavonoid, lacton sesquiterpene,
    ↳ anthocyan, và tannin. Những hợp chất này có tác dụng được lý quan trọng, hỗ trợ chức
    ↳ năng gan và sức khỏe tổng thể.",
  "knowledge": "2. Thành phần hóa học chứa trong Actiso"
},
{
  "question": "Tác dụng lợi tiểu và lợi mật của Actiso được thể hiện như thế nào trong việc điều
    ↳ trị bệnh phù, thấp khớp và các vấn đề liên quan đến gan?",
  "answer": "Actiso có tác dụng lợi tiểu, giúp thông tiểu tiện và thông mật, từ đó hỗ trợ điều trị
    ↳ bệnh phù, thấp khớp, và các bệnh yếu gan thận.",
  "knowledge": "3. Công dụng của Actiso"
},
{
  "question": "Cách sử dụng và liều dùng Actiso như thế nào để đạt hiệu quả tốt nhất trong điều
    ↳ trị và bảo vệ sức khỏe?",
  "answer": "Actiso có thể dùng lá tươi hoặc khô, thuốc sắc từ 5-10% hoặc 2-10g lá khô/ngày. Có
    ↳ thể chế thành cao mềm hay lỏng để sử dụng dưới dạng giọt, mỗi lần 10-40 giọt.",
  "knowledge": "4. Cách dùng và liều dùng Actiso"
},
{
  "question": "Có những nghiên cứu nào đã chứng minh về tác dụng của Actiso trong việc bảo vệ
    ↳ gan, hỗ trợ tiêu hóa, và ngăn ngừa các bệnh lý khác không?",
  "answer": "Nhiều nghiên cứu đã chứng minh Actiso giúp bảo vệ gan, tăng bài tiết dịch mật, hỗ
    ↳ trợ tiêu hóa và ngăn ngừa các bệnh lý như tim mạch và ung thư.",
  "knowledge": "5. Các nghiên cứu về tác dụng của Actiso"
},
{
  "question": "Liệu Actiso có thể được sử dụng trong việc hỗ trợ điều trị ung thư và giảm nguy cơ
    ↳ mắc bệnh không? Cơ chế hoạt động của nó là gì?",
  "answer": "Actiso chứa các chất chống oxy hóa có thể giúp ngăn ngừa tổn thương tế bào và biến
    ↳ đổi gene dẫn đến ung thư, do đó có thể giảm nguy cơ mắc bệnh.",
  "knowledge": "5. Các nghiên cứu về tác dụng của Actiso"
},
{
  "question": "Trong y học cổ truyền, có những lời khuyên cụ thể nào về cách sử dụng và kết hợp
    ↳ Actiso với các loại thực phẩm khác để tối ưu hóa tác dụng của nó?",
  "answer": "Để tối ưu hóa tác dụng của Actiso, nên kết hợp với chế độ ăn uống cân bằng, lành
    ↳ mạnh và có thể sử dụng cùng với các loại thực phẩm giàu chất xơ để hỗ trợ tiêu hóa.",
  "knowledge": "3. Công dụng của Actiso"
},
{
  "question": "Có những tác dụng phụ nào có thể xảy ra khi sử dụng Actiso trong điều trị dài
    ↳ hạn? Làm thế nào để đối phó với những tác dụng này?",
  "answer": "Actiso không gây độc, nhưng nếu có tác dụng phụ, cần theo dõi sức khỏe và tham
    ↳ khảo ý kiến bác sĩ để điều chỉnh liều dùng hoặc ngừng sử dụng.",
  "knowledge": "4. Cách dùng và liều dùng Actiso"
}
]

```

Figure 8: Sample output of the Answer Generation step for the Entire Content .

```

[
  {
    "question": "How does Artichoke help protect the liver and detoxify it? What is its mechanism  

    ↳ of action?",
    "answer": "Artichoke contains cynarin and caffeoylquinic acid, which are antioxidants that can  

    ↳ protect the liver. They help reduce levels of harmful toxins in the liver and may assist in  

    ↳ repairing damaged liver cells.",
    "knowledge": "5. Studies on the effects of Artichoke."
  },
  {
    "question": "What are the main chemical components of Artichoke, and how do these  

    ↳ components affect liver health and overall health?",
    "answer": "Artichoke contains compounds such as caffeic acid derivatives, flavonoids,  

    ↳ sesquiterpene lactones, anthocyanins, and tannins. These compounds have important  

    ↳ pharmacological effects, supporting liver function and overall health.",
    "knowledge": "2. The chemical composition contained in Artichoke."
  },
  {
    "question": "How are the diuretic and choleretic effects of Artichoke manifested in the treatment  

    ↳ of edema, rheumatism, and liver-related issues?",
    "answer": "Artichoke has diuretic effects, helping with urination and bile secretion, thereby  

    ↳ supporting the treatment of edema, rheumatism, and weak liver and kidney diseases.",
    "knowledge": "3. Uses of Artichoke"
  },
  {
    "question": "How should Artichoke be used, and what is the recommended dosage to achieve the  

    ↳ best results in treatment and health protection?",
    "answer": "Artichoke can be used fresh or dried, decocted at 5-10%, or 2-10g of dried leaves per  

    ↳ day. It can be prepared as a soft or liquid extract and used in drops, 10-40 drops per dose.",
    "knowledge": "4. How to use and dosage of Artichoke"
  },
  {
    "question": "Are there any studies that have proven the effects of Artichoke in protecting the  

    ↳ liver, aiding digestion, and preventing other diseases?",
    "answer": "Many studies have proven that Artichoke helps protect the liver, increase bile  

    ↳ secretion, aid digestion, and prevent diseases such as cardiovascular disease and cancer.",
    "knowledge": "5. Studies on the effects of Artichoke"
  },
  {
    "question": "Can Artichoke be used to support cancer treatment and reduce the risk of disease?  

    ↳ What is its mechanism of action?",
    "answer": "Artichoke contains antioxidants that may help prevent cell damage and genetic  

    ↳ mutations leading to cancer, thereby potentially reducing the risk of disease.",
    "knowledge": "5. Studies on the effects of Artichoke"
  },
  {
    "question": "In traditional medicine, are there specific recommendations on how to use and  

    ↳ combine Artichoke with other foods to optimize its effects?",
    "answer": "To optimize the effects of Artichoke, it should be combined with a balanced, healthy  

    ↳ diet and can be used alongside fiber-rich foods to support digestion.",
    "knowledge": "3. Uses of Artichoke"
  },
  {
    "question": "What side effects might occur when using Artichoke for long-term treatment? How  

    ↳ can these side effects be managed?",
    "answer": "Artichoke is non-toxic, but if side effects occur, monitor health and consult a doctor  

    ↳ to adjust the dosage or discontinue use.",
    "knowledge": "4. How to use and dosage of Artichoke"
  }
]

```

Figure 9: Sample output of the Answer Generation step for the Entire Content (translated to English).

```

[
  {
    "question": "Tôi muốn hỏi về phương pháp truyền thống dùng dược liệu để điều trị bệnh sổ mũi  

    → và say nắng, bạn có thể chia sẻ thông tin về điều này không?" ("I would like to ask about  

    → traditional methods using medicinal herbs to treat a runny nose and heatstroke; can you  

    → share information about this?"),
    "answer": "Để điều trị bệnh sổ mũi và say nắng, bạn có thể sử dụng 12g thân cây Bướm bạc, 3g  

    → Bạc hà, và 10g lá Ngũ trảo. Rửa sạch và để ráo nước, sau đó đun sôi cùng với nước để  

    → dùng thay thế cho nước trà hằng ngày." ("To treat a runny nose and heatstroke, you can  

    → use 12g of Butterfly tree stem, 3g of Mint, and 10g of Vitex leaves. Clean and drain, then  

    → boil with water to use as a daily tea substitute."),
    "knowledge": "4. Một số bài thuốc kinh nghiệm" ("4. Some empirical remedies")
  },
  {
    "question": "Có bài thuốc dân gian nào được sử dụng trong y học cổ truyền để chữa bệnh ho, sốt  

    → và sưng amidan không?" ("Is there a folk remedy used in traditional medicine to treat  

    → cough, fever, and tonsillitis?"),
    "answer": "Có, để chữa bệnh ho, sốt và sưng amidan, bạn có thể sử dụng 30g rễ cây Bướm bạc,  

    → 10g rễ Bọ mẩy và 20g Huyền sâm. Rửa sạch các vị thuốc, sau đó sắc với một lượng nước  

    → phù hợp và sử dụng." ("Yes, to treat cough, fever, and tonsillitis, you can use 30g of  

    → Butterfly tree root, 10g of Mẩy root, and 20g of Scrophularia root. Clean the herbs, then  

    → decoct with an appropriate amount of water and use."),
    "knowledge": "4. Một số bài thuốc kinh nghiệm" ("4. Some empirical remedies")
  },
  {
    "question": "Xin bác sĩ tư vấn về việc sử dụng dược liệu từ cây Bướm bạc để điều trị viêm thận,  

    → phù và giúp lợi tiểu." ("Please advise on using medicinal herbs from the Butterfly tree to  

    → treat nephritis, edema, and help with diuresis."),
    "answer": "Để điều trị viêm thận, phù và giúp lợi tiểu, bạn có thể sử dụng 30g thân Bướm bạc  

    → kết hợp với 60g Kim ngân hoa và 30g Mã đề, sắc nước uống." ("To treat nephritis, edema,  

    → and help with diuresis, you can use 30g of Butterfly tree stem combined with 60g of  

    → Honeysuckle flower and 30g of Plantain, and decoct for drinking."),
    "knowledge": "4. Một số bài thuốc kinh nghiệm" ("4. Some empirical remedies")
  },
  {
    "question": "Tôi đang quan tâm đến bài thuốc truyền thống dành cho bệnh khí hư bạch đới, bạn  

    → có thể chia sẻ thông tin chi tiết về cách sử dụng không?" ("I am interested in a traditional  

    → remedy for leucorrhea, can you share detailed information on how to use it?"),
    "answer": "Để chữa bệnh khí hư bạch đới, bạn cần 10 đến 20g rễ Bướm bạc. Rửa sạch, sắc kỹ với  

    → nước lọc một lượng phù hợp và sử dụng mỗi ngày." ("To treat leucorrhea, you need 10 to  

    → 20g of Butterfly tree root. Clean it thoroughly, decoct with a suitable amount of filtered  

    → water, and use daily."),
    "knowledge": "4. Một số bài thuốc kinh nghiệm" ("4. Some empirical remedies")
  },
  {
    "question": "Bác sĩ có thể giới thiệu về cách sử dụng dược liệu từ lá cây Mướp và lá cây Bướm  

    → bạc để chữa lở loét da không?" ("Can you introduce how to use medicinal herbs from  

    → Loofah leaves and Butterfly tree leaves to treat skin ulcers?"),
    "answer": "Để chữa lở loét da, bạn cần dùng lá cây Mướp tươi và lá cây Bướm bạc tươi với liều  

    → lượng bằng nhau. Rửa sạch, để ráo rồi giã nát ra đắp bã vào các vùng lở loét da, cố định  

    → lại và sau đó rửa lại với nước sạch." ("To treat skin ulcers, you need fresh Loofah leaves  

    → and fresh Butterfly tree leaves in equal amounts. Clean them, drain, then crush and apply  

    → the paste to the ulcerated areas, secure it, and then wash with clean water."),
    "knowledge": "4. Một số bài thuốc kinh nghiệm" ("4. Some empirical remedies")
  }
]

```

Figure 10: Sample output of the Answer Generation step for the Herbal Prescriptions Content.