

**The 11th Asian-Australasian Conference on Precision Agriculture (ACPA 11)
October 14-16, 2025, Chiayi, Taiwan**

Applying Retrieval-Augmented-Generation to Support Farmers in Pest and Disease Diagnosis

Sheng-Chieh Lai ¹, Yan-Fu Kuo ^{1*}

¹ Department of Biomechatronics Engineering, National Taiwan University, Taipei, Taiwan

*Corresponding Author: ykuo@ntu.edu.tw

ABSTRACT

According to the Ministry of Agriculture, crop production in Taiwan reached a value of \$275 billion NTD in 2023, highlighting the economic importance of agriculture. However, the industry is now facing serious challenges, particularly in pest and disease identification and crop protection. Due to global ecological challenges, the manifestations of local pests and diseases have changed, making it difficult for farmers to rely on past experiences to identify and manage them effectively. Farmers now require more assistance from agricultural experts. However, the overwhelming number of inquiries overburden these experts, leading to delays and inefficiencies. Therefore, there is a pressing need for an automated system to assist farmers with pest and disease management. This study aims to apply retrieval-augmented-generation and develop a smart farming question-answering system to provide farmers with accurate and practical guidance on pest and disease management. The agricultural knowledge documents were collected from the organizations under the Ministry of Agriculture and stored in a database. Given a query, the hybrid retriever would perform the initial selection from the database, which leverages both sparse and dense retrievals in a parallel setup to capture complementary semantic and lexical signals. The retrieved candidates are then further re-ranked using a reranker model, which scores the relevance between the query and each candidate more precisely, enhancing the overall retrieval quality. Finally, the top-ranked candidates are passed to a large language model, which generates a Chinese response based on the retrieved information. The two-stage retriever achieved a top-1 accuracy of 97.4% on Delta Reading Comprehension Dataset, which consist of 1000 unique passages and 3493 corresponding questions. The result showed that the proposed approach can assist in pest and disease management while reducing the labor and time required.

Keywords: Pest and Disease Identification, Retrieval-Augmented Generation (RAG), Large Language Model (LLM)

INTRODUCTION

Due to global climate change and shifting ecological conditions, the types and manifestations of local pests and diseases have become more diverse and unpredictable, making it difficult for farmers to rely solely on past experiences for effective management. This situation has increased the demand for professional guidance from agricultural experts. Yet, the large volume of inquiries often overwhelms these experts, resulting in delays, inefficiencies. Consequently, there is a pressing need for a question-answering system that can assist

farmers in accurately identifying pests and diseases, while also providing timely and practical crop protection strategies.

MATERIALS AND METHODS

2.1 Data Collection and Preprocessing

The text data was obtained in PDF format from 10 organizations under the Ministry of Agriculture, Taiwan. At this stage, only publications from the past 10 years have been collected. The data was manually preprocessed by removing all images and tables, while retaining the remaining textual content. The cleaned text was subsequently converted into JSON format.

2.2 RAG PIPELINE

The pipeline consists of a two-stage retriever and an LLM. In the first stage, a hybrid retriever (Bormotov, 2024) is used, consisting of a BM25 module and an embedding model (jinaai/jina-embeddings-v3) running in parallel. Their results are combined to select the top-N most relevant candidates. In the second stage, a reranker model (BAAI/bge-reranker-v2-m3) takes these top-N candidates and reorders them, then selects the top-K most relevant candidates. Finally, the top-ranked candidates are passed to an LLM (Breeze-7B-Instruct-v1_0), which generates a Chinese response based on the retrieved information.

RESULTS & DISCUSSION

The retriever was evaluated on the Delta Reading Comprehension Dataset (DRCD), which contains 1,000 unique passages and 3,493 corresponding questions. The results are showed in Table 1. The two-stage retriever achieved a top-1 accuracy of 97.4% and a top-5 accuracy of 99.7%, with an average response time of only 0.3 seconds. These results demonstrate that the two-stage retriever can achieve high accuracy while maintaining low latency.

Table 1 Two-stage retriever accuracy and speed.

Type	Top-1 accuracy	Top-5 accuracy	Average response time (sec)
Two-stage retriever (N=50, K=5)	97.4%	99.7%	0.265
Reranker: directly select Top-5	97.5%	99.7%	3.93
BM25 + embedding: directly select Top-5	92.3%	99.1%	0.033

CONCLUSIONS

We constructed a RAG-based question-answering system, integrating domain-specific documents with an LLM. Using a two-stage retriever (BM25, embedding model, and reranker), the system achieved 99.7% top-5 accuracy with an average response time of 0.3 seconds on the DRCD dataset. These results demonstrate the system's potential to assist in pest and disease management while reducing the labor and time required.

ACKNOWLEDGEMENTS

We gratefully acknowledge the Taiwan Agricultural Research Institute, as well as other organizations under the Ministry of Agriculture, Taiwan, for their assistance and support in this study.

REFERENCES

Bormotov, K. (2024). Hybrid retrieval: Combining Bert and BM25 for enhanced performance. Retrieved from <https://medium.com/@bormotovk/hybrid-retrieval-combining-bert-and-bm25-for-enhanced-performance-4f6f80881c13>