ISSN 2277-2685
IJESR/April. 2026/ Vol-16/Issue-2s1/303-307
Sarfaraz khan *et. al.*, / International Journal of Engineering & Science Research

Full Length Article

Enhancing Agricultural Decision-Making with Combined Language Models

Sarfaraz khan[1],Syed asadullah[2],Abdul ahmed khan[3], Dr.S.Md.Mazhar ul haq[4]
[1.2.3]B.E.Students; Department of Artificial Intelligence and Data Science ISL Engineering College Hyderabad, Telangana India
[4]Associate professor; Department of Artificial Intelligence and Data Science ISL Engineering College Hyderabad, Telangana India
Email;Sarfraxy@gmail.com,Syedasadullah571@gmail.com,abdulahmedkhan04@gmail.com
Accepted 27-04-2026

ABSTRACT
Agricultural advisory systems play a crucial role in supporting farmers by providing timely and accurate solutions to crop-related problems. However, building reliable question-answering systems for agriculture is challenging due to the diverse, domain-specific, and often noisy nature of farmer queries. In this project, a retrieval-based agricultural question-answering framework is proposed to deliver accurate and trustworthy responses by leveraging historical expert-curated data. The system utilizes a pretrained sentence embedding model to convert farmer queries into dense vector representations, followed by similarity-based retrieval using FAISS to identify the most relevant past questions and their corresponding expert answers. Unlike generative models, which often produce vague or hallucinated responses, the proposed approach ensures factual correctness by reusing validated advisory content. Experiments conducted on a large farmer query dataset demonstrate that the system achieves a training accuracy of 100% and a test accuracy of 95.9%, with qualitative evaluation confirming high relevance and agricultural validity of the retrieved responses. The results indicate that retrieval-based question answering is a practical, efficient, and reliable solution for real-world agricultural advisory applications.

Keywords— *Agricultural Advisory System, Retrieval-Based Question Answering, Sentence Embeddings, FAISS, Natural Language Processing (NLP), Information Retrieval, Semantic Similarity, Farmer Query Analysis, Expert Advisory Retrieval, Dense Vector Representation, Pretrained Language Models, Agricultural Informatics.*

Introduction

Agriculture is a vital sector that directly influences food security, economic stability, and rural livelihoods. Farmers frequently face challenges related to crop diseases, pest infestations, nutrient deficiencies, and government agricultural schemes, and timely access to accurate advisory information is crucial for effective decision-making. With the rapid growth of digital platforms, large volumes of farmer queries are generated through call centers, mobile applications, and online portals. However, providing reliable, consistent, and domain-specific answers to these queries remains a significant challenge due to the diversity and complexity of agricultural problems. Recent advancements in natural language processing and machine learning have enabled the development of automated agricultural question-answering systems. While generative large language models have shown promise in understanding natural language, they often suffer from issues such as hallucination, inconsistency, and high computational cost when applied to real-world advisory systems. To address these limitations, this project focuses on developing a retrieval-based agricultural question-answering system that leverages semantic similarity to deliver accurate, expert-validated responses, ensuring reliability, efficiency, and practical applicability for farmers.

LITERATURE REVIEW

Title: Research on the Training and Application Methods of a Lightweight Agricultural Domain-Specific Large Language Model Supporting Mandarin Chinese and Uyghur
Author(s): K. Pan, X. Zhang, and L. Chen
Year: 2024

Description: This study focuses on the development of a lightweight domain-specific large language model tailored for agricultural applications, particularly supporting Mandarin Chinese and Uyghur languages. The authors propose optimized training strategies to reduce computational complexity while maintaining acceptable performance in agricultural text understanding and generation tasks. The work highlights the importance of domain adaptation and multilingual support in agricultural advisory systems. However, the model primarily relies on generative responses, which may still suffer from hallucination and limited reliability in real-world advisory scenarios.

Title: Impact of Model Size on Fine-Tuned LLM Performance in Data-to-Text Generation: A State-of-the-Art Investigation

Author(s): J. Mahapatra and U. Garain

Year: 2024

Description: This paper investigates how the size of large language models affects their performance when fine-tuned for data-to-text generation tasks. The authors analyze multiple LLM variants and demonstrate that larger models generally achieve better fluency and coherence, but at the cost of higher computational requirements. The study emphasizes that increasing model size does not always guarantee proportional performance gains, especially in domain-specific applications, making lightweight or alternative approaches more practical for real-world deployments.

Title: LLM-Based Text Prediction and Question Answer Models for Aphasia Speech

Author(s): S. B. Manir, K. M. S. Islam, P. Madiraju, and P. Deshpande

Year: 2024

Description: This work explores the use of large language models for text prediction and question-answering to assist individuals with aphasia. The authors demonstrate that LLM-based systems can generate meaningful responses when fine-tuned on task-specific data. Although the domain differs from agriculture, the study highlights key challenges associated with generative models, such as response inconsistency and dependence on high-quality training data, which are also relevant to agricultural advisory systems.

Title: ChatAgri: Exploring Potentials of ChatGPT on Cross-Linguistic Agricultural Text Classification

Author(s): B. Zhao, W. Jin, J. D. Ser, and G. Yang

Year: 2023

Description: The ChatAgri study evaluates the effectiveness of ChatGPT for agricultural text classification across multiple languages. The authors show that large pretrained language models can generalize well across linguistic boundaries with minimal fine-tuning. However, the study also points out that ChatGPT-based systems may generate generalized or ambiguous outputs, limiting their reliability for precise agricultural decision-making tasks such as pest control and fertilizer recommendations.

METHODOLOGIES

MODULES NAME:

- **Data Collection and Preprocessing Module**
- **Sentence Embedding Generation Module**
- **Similarity Search and Retrieval Module**
- **Answer Selection Module**
- **Evaluation and Performance Analysis Module**
- **Model Storage and Deployment Module**

MODULES EXPLANATION:

1. Data Collection and Preprocessing Module: This module is responsible for loading the agricultural question–answer datasets, including training, validation, and testing data. The data is cleaned by removing duplicates, handling missing values, and normalizing text through lowercasing and trimming. Proper preprocessing ensures that the input data is consistent and suitable for semantic representation, which directly impacts retrieval accuracy.

2. Sentence Embedding Generation Module: In this module, farmer questions are converted into dense numerical vectors using a pretrained sentence transformer model. These embeddings capture the semantic meaning of the text rather than relying on exact keyword matching. By representing queries in a high-dimensional embedding space, the system can effectively identify semantically similar questions even when wording differs.

3. Similarity Search and Retrieval Module: This module performs fast similarity search using a vector indexing mechanism based on cosine similarity. A FAISS index is constructed using the embeddings of training questions. When a new query is received, its embedding is compared against the indexed vectors to retrieve the most similar historical question. This enables efficient and accurate retrieval even for large datasets.

4. Answer Selection Module: Once the most relevant historical question is identified, this module retrieves its corresponding expert-validated answer. Since the answers are sourced from previously verified advisory data, the system ensures factual correctness and consistency. This module eliminates the risk of hallucination by avoiding answer generation.

5. Evaluation and Performance Analysis Module: This module evaluates the effectiveness of the system using retrieval accuracy metrics based on similarity thresholds. Both training and testing datasets are used to assess system performance. Additionally, qualitative analysis is performed by comparing retrieved answers with ground truth responses to ensure agricultural relevance and correctness.

6. Model Storage and Deployment Module: The final module handles saving the trained sentence embedding model, FAISS index, and processed datasets for future use. This enables easy loading and deployment of the system in real-time applications such as web portals or mobile advisory platforms. The modular design allows seamless updates when new agricultural data becomes available.

SOFTWARE TESTING

Types of Tests

Unit Testing

Unit testing involves the design of test cases that validate whether the internal program logic is functioning properly and whether program inputs produce valid outputs. All decision branches and internal code flow are carefully validated to ensure correctness. It is the testing of individual software units of the application and is performed after the completion of each module before integration with other modules. This is a structural testing method that relies on knowledge of the internal construction of the system and is invasive in nature.

Unit tests perform basic testing at the component level and evaluate specific business processes, application modules, and system configurations independently. Unit testing ensures that each unique path of the process performs accurately according to the documented specifications and contains clearly defined inputs and expected outputs. It also helps in identifying coding errors, logic faults, and inconsistencies at an early stage of development, thereby reducing the overall debugging effort during integration.

In the proposed YOLOv10-based surgical instrument detection system, unit testing was performed on modules such as image preprocessing, dataset loading, feature extraction, hybrid attention mechanisms, object detection heads, and result visualization. Each module was tested independently to verify that the outputs generated were correct and reliable before combining them into the complete system.

Functional Test

Functional testing provides systematic demonstrations that the functions tested are available and operate according to the business requirements, technical requirements, system documentation, and user manuals. This testing focuses on validating the functional behavior of the software rather than its internal implementation.

Functional testing is centered on the following items:

Valid Input: Identified classes of valid input must be accepted and processed correctly.

Invalid Input: Identified classes of invalid input must be rejected appropriately without system failure.

Functions: Identified functions and operations must be exercised thoroughly.

Output: Identified classes of application outputs must be verified for correctness and accuracy.

Systems/Procedures: Interfacing systems or procedures must be properly invoked and executed.

In this project, functional testing was conducted to verify operations such as uploading laparoscopic images and videos, preprocessing inputs, detecting surgical instruments, displaying prediction results with bounding boxes, and generating evaluation metrics correctly. The Flask-based web application interface was also tested to ensure proper user interaction, smooth navigation, and accurate output generation.

System Test

System testing ensures that the entire integrated software system meets the specified functional and non-functional requirements. It tests the complete configuration of the system to ensure known and predictable results under different conditions. System testing is based on process descriptions and workflows, emphasizing integration points and interactions among different modules.

The main objective of system testing is to validate the overall behavior of the application after integrating all individual components. It checks whether the software system works as expected in a real-time environment and verifies end-to-end functionality.

In the YOLOv10 surgical instrument detection system, system testing was carried out to ensure smooth integration between image preprocessing, feature extraction, hybrid attention modules, YOLOv10 model inference, Flask web application, and visualization modules. The testing confirmed that the complete system performs efficiently for real-time surgical instrument detection and provides accurate outputs without failures.

Performance Test

Performance testing ensures that outputs are produced within acceptable time limits and evaluates the time taken by the system for processing requests, compiling results, and responding to user interactions. This testing is important for determining the stability, scalability, and responsiveness of the software under different workloads.

Performance testing measures various parameters such as execution speed, response time, throughput, memory usage, CPU utilization, and overall system efficiency. The primary goal is to ensure that the application performs reliably even when processing large datasets or real-time video streams.

In this project, performance testing was conducted to analyze metrics such as inference speed, Frames Per Second (FPS), response time, memory utilization, and real-time detection efficiency. The YOLOv10 model was evaluated on laparoscopic videos to ensure that surgical instruments were detected accurately while maintaining fast processing speed suitable for real-time surgical assistance applications.

FUTURE ENHANCEMENT

The proposed retrieval-based agricultural question-answering system can be further enhanced in several ways to improve its functionality and real-world applicability. Future work may include the integration of multilingual support to handle farmer queries in regional languages, enabling wider accessibility. The system can also be extended to incorporate region-specific and crop-specific filtering to provide more personalized advisory responses. Additionally, combining retrieval with controlled generative models in a hybrid framework can help generate more detailed explanations while maintaining factual correctness. The inclusion of real-time data sources such as weather forecasts, soil health records, and market prices can further improve decision support for farmers. Finally, deploying the system as a mobile application or integrating it with government agricultural helplines and chatbot platforms can significantly enhance its impact and usability in large-scale agricultural advisory services.

CONCLUSION

In this project, a reliable and efficient agricultural question-answering system was successfully developed using a retrieval-based approach. By leveraging semantic sentence embeddings and similarity-based search techniques, the system provides accurate and consistent advisory responses by retrieving expert-validated answers from historical data. The proposed method effectively overcomes the limitations of generative large language models, such as hallucination, inconsistency, and high computational cost. Experimental results demonstrate that the system achieves high performance, with a test accuracy of approximately 95.9%, and qualitative evaluation confirms the agricultural relevance and correctness of the retrieved responses. The modular design, low computational requirements, and high reliability make the proposed system well-suited for real-world agricultural advisory applications, contributing to improved decision-making and support for farmers.

REFERENCES

[1] Y. Chang, X. Wang, J. Wang, Y. Wu, L. Yang, K. Zhu, H. Chen, X. Yi,C. Wang, Y. Wang, W. Ye, Y. Zhang, Y. Chang, P. S. Yu, Q. Yang, and X. Xie, ‘‘A survey on evaluation of large language models,’’ ACM Trans.Intell. Syst. Technol., vol. 15, no. 3, pp. 1–45, Mar. 2024.

[2] J. Howard and S. Ruder, ‘‘Universal language model fine-tuning for textclassification,’’ 2018, arXiv:1801.06146.

[3] Y. Bengio, R. Ducharme, and P. Vincent, ‘‘A neural probabilistic language model,’’ in Proc. Adv. Neural Inf. Process. Syst., vol. 13, Mar. 2000,pp. 1–11.

[4] K. Tan, ‘‘Large language models for crop yield prediction,’’ Tech. Rep.,2024.

[5] Y. Luo, X. Cai, J. Qi, D. Guo, and W. Che, ‘‘FPGA—Accelerated CNN for real-time plant disease identification,’’ Comput. Electron. Agricult.,vol. 207, May 2023, Art. no. 107715.

[6] Y. Xu, J. Lu, and J. Zhang, ‘‘Bridging the gap between different vocabularies for LLM ensemble,’’ 2024, arXiv:2404.09492.

[7] T. Wang, X. Xu, C. Wang, Z. Li, and D. Li, ‘‘From smart farming towards unmanned farms: A new mode of agricultural production,’’ Agriculture, vol. 11, no. 2, p. 145, Feb. 2021.

[8] J.Li,M.Xu,L.Xiang,D.Chen,W.Zhuang,X.Yin,andZ.Li,‘‘Foundation models in smart agriculture: Basics, opportunities, and challenges,’’Comput. Electron. Agricult., vol. 222, Jul. 2024, Art. no. 109032.

[9] K. Pan, X. Zhang, and L. Chen, ‘‘Research on the training and applicationmethods of a lightweight agricultural domain-specific large languagemodel supporting Mandarin Chinese and Uyghur,’’ Appl. Sci., vol. 14,no. 13, p. 5764, Jul. 2024.

[10] J. Mahapatra and U. Garain, ‘‘Impact of model size on fine-tuned LLMperformance in data-to-text generation: A state-of-the-art investigation,’’2024, arXiv:2407.14088.

[11] S. B. Manir, K. M. S. Islam, P. Madiraju, and P. Deshpande, ‘‘LLM-basedtext prediction and question answer models for aphasia speech,’’ IEEEAccess, vol. 12, pp. 114670–114680, 2024.

[12] B. Zhao, W. Jin, J. D. Ser, and G. Yang, ‘‘ChatAgri: Exploringpotentials of ChatGPT on cross-linguistic agricultural text

classification,''Neurocomputing, vol. 557, Nov. 2023, Art. no. 126708.
[13] J. Wang, ''LLM-based fine-tuning data generation for relation tripletextraction with expert ensemble and demonstration selection,'' in Proc.IEEE 12th Int. Conf. Intell. Syst. (IS), Aug. 2024, pp. 1–7.
[14] M. P. Geetha and D. K. Renuka, ''Improving the performance of aspectbased sentiment analysis using fine-tuned BERTbaseuncasedmodel,''Int.J. Intell. Netw., vol. 2, pp. 64–69, Mar. 2021.
[15] H. Wang, H. Wu, H. Zhu, Y. Miao, Q. Wang, S. Qiao, H. Zhao, C. Chen,and J. Zhang, ''A residual LSTM and Seq2Seq neural network based onGPT for Chinese rice-related question and answer system,'' Agriculture,vol. 12, no. 6, p. 813, Jun. 2022.
[16] S. Rezayi, Z. Liu, Z. Wu, C. Dhakal, B. Ge, Z. Chen, T. Liu, andS. Li, ''AgriBERT: Knowledge-infused agricultural language models formatching food and nutrition,'' in Proc. IJCAI, Jul. 2022, pp. 5150–5156.
[17] J. Qing, X. Deng, Y. Lan, and Z. Li, ''GPT-aided diagnosis on agriculturalimage based on a new light YOLOPC,'' Comput. Electron. Agricult.,vol. 213, Oct. 2023, Art. no. 108168.
[18] D. Allemang and J. Sequeda, ''Increasing the LLM accuracy for questionanswering: Ontologies to the rescue!'' 2024, arXiv:2405.11706.
[19] S. Rezayi, Z. Liu, Z. Wu, C. Dhakal, B. Ge, H. Dai, G. Mai, N. Liu,C. Zhen, T. Liu, and S. Li, ''Exploring new frontiers in agriculturalNLP: Investigating the potential of large language models for foodapplications,'' IEEE Trans. Big Data, early access, Aug. 15, 2024, doi:10.1109/TBDATA.2024.3442542.
[20] J. Kpodo, P. Kordjamshidi, and A. P. Nejadhashemi, ''AgXQA: Abenchmark for advanced agricultural extension question answering,''Comput. Electron. Agricult., vol. 225, Oct. 2024, Art. no. 109349.
[21] X. Guo, J. Wang, G. Gao, J. Zhou, Y. Li, Z. Cheng, and G. Miao,''Efficient agricultural question classification with a BERT-enhancedDPCNNmodel,'' IEEE Access, vol. 12, pp. 109255–109268, 2024.
[22] S.Yang,Z.Yuan,S.Li, R.Peng, K.Liu, andP. Yang,''GPT-4 asevaluator:Evaluating large language models on pest management in agriculture,''2024, arXiv:2403.11858.