AgLLM: Enhancing Agriculture with LLM-Driven Data Curation and Verifiable Recommendations

Muhammad Arbab Arshad, Hossein Zaremehrjerdi, Talukder Zaki Jubery, Tirtho Roy, Rim Nassiri, Asheesh K. Singh, Arti Singh, Chinmay Hegde, Baskar Ganapathysubramanian, Aditya Balu, Adarsh Krishnamurthy, Soumik Sarkar(soumiks@iastate.edu)

Iowa State University

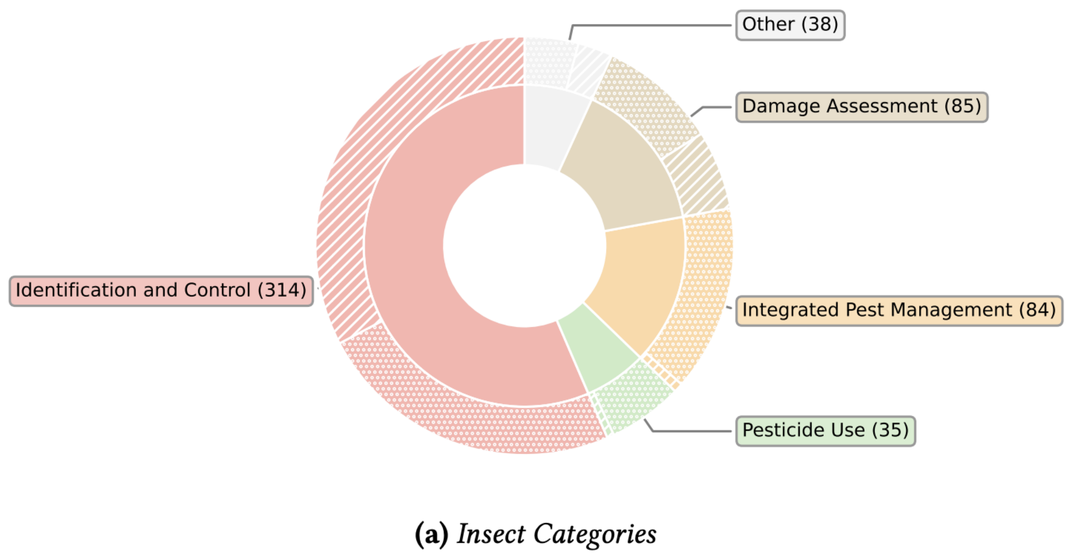
Motivation

Purpose

- Consolidate fragmented agricultural insect/weed data
- Provide verifiable recommendations for improved pest management

Data Sources

- Expert-verified data on 90 species (49 insects, 41 weeds)
- Documents from University Extension Programs



(a) *Insect Categories*

Methodology

Knowledge Base Construction:

- Academic literature and curated internet resources

Retrieval Augmented Generation (RAG) System:

- Document pre-processing, indexing, and metadata-based filtering

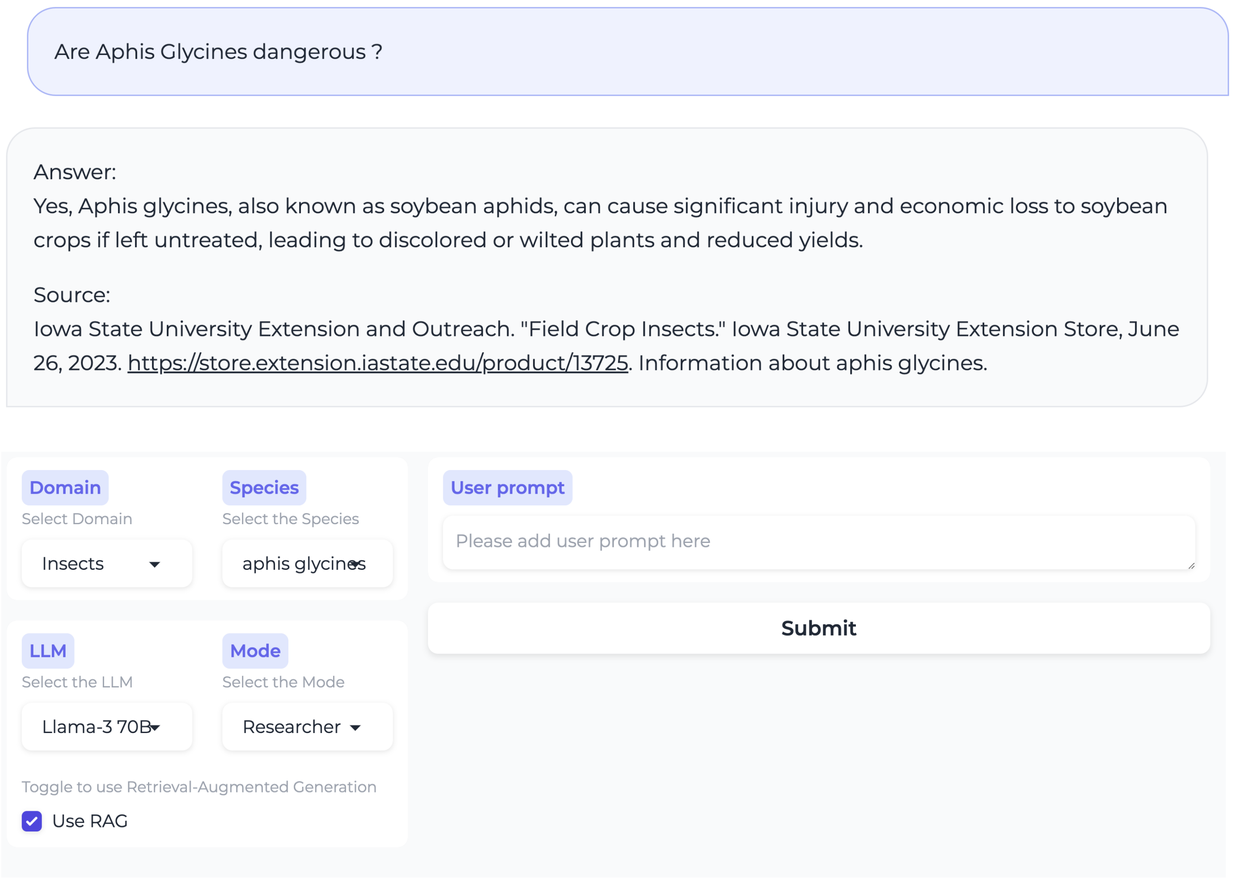
Conversational Interface:

- ConversationalRetrievalChain (Langchain) with Gradio UI

Multi-LLM Support:

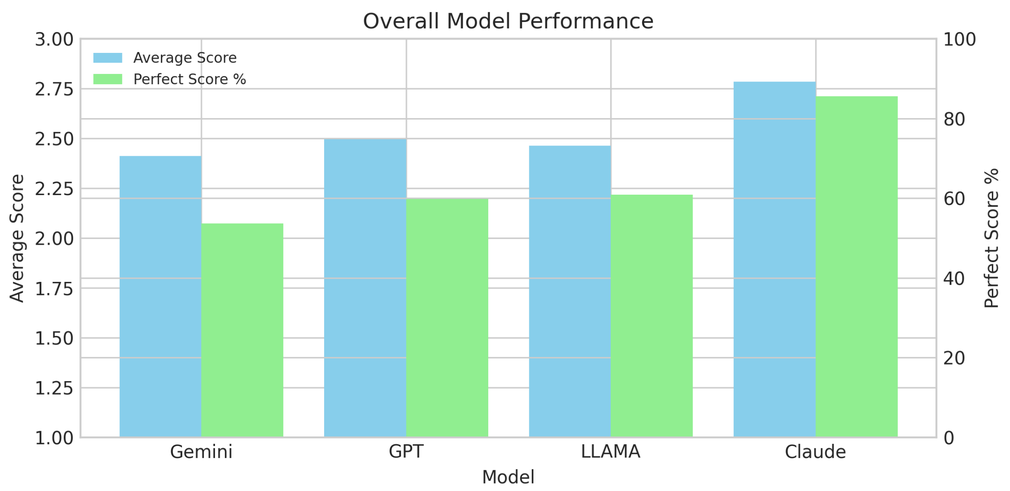
- GPT-4, GPT-3.5, Claude 3 Opus, Llama-3, Gemini-1.5 Pro

- Synthetic question-answer pair generation for Evaluation



Results

- Expert assessment of LLM outputs



- RAG pipeline evaluation:

  Recall of 82% (at n=3) and Precision 65% (at n=1)

Conclusion

- Developed RAG system with high recall (82% at n=3) and precision (65% at n=1)
- Implemented user-type aware responses for farmers and researchers
- Bridged gap between scattered information and practical pest management needs

National Institute of Food and Agriculture
U.S. DEPARTMENT OF AGRICULTURE
IOWA STATE UNIVERSITY
OF SCIENCE AND TECHNOLOGY

AI Institute for Resilient Agriculture (AIIRA) is supported by the National Science Foundation (NSF) and United States Department of Agriculture - National Institute of Food and Agriculture (award #2021-67021-35329)