

Fine-tuned LLM with RAG for Tennessee Agricultural Extension Applications

McKensie Nelms and Jonathan Yoder
Department of Biosystems Engineering and Soil Science

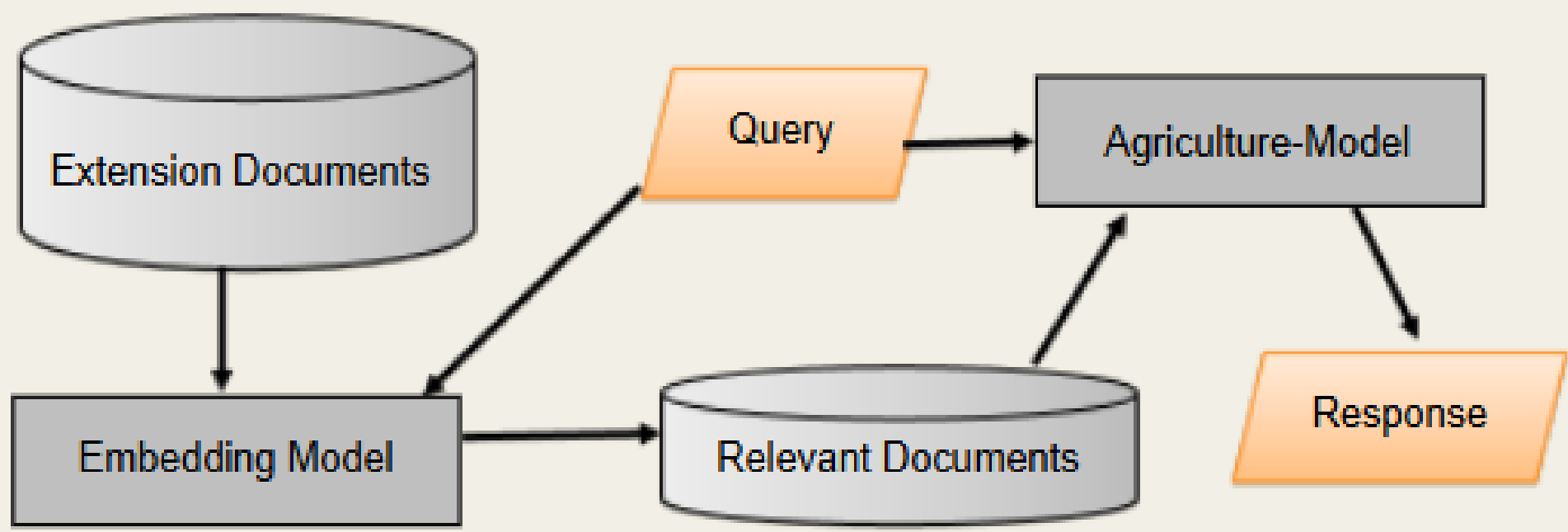
Motivation

Agricultural Extension work is responsible for sharing agricultural research and knowledge with farmers and producers. While extension agents are available, it would be convenient to have a tool which can aid both agents and stakeholders. This work aims to create an extension tool for Tennessee which can answer questions and disseminate knowledge in a reliable format. This tool can be adapted to other regions by changing the RAG database with regional extension sources.

Main Research Objectives:

1. Fine-tune large language model (LLM) on general agricultural knowledge
2. Create a pipeline integrating the agriculture-informed model with retrieval augmented generation (RAG) using Tennessee extension resources as data
3. Evaluate results for reliability and legibility

Retrieval Augmented Generation

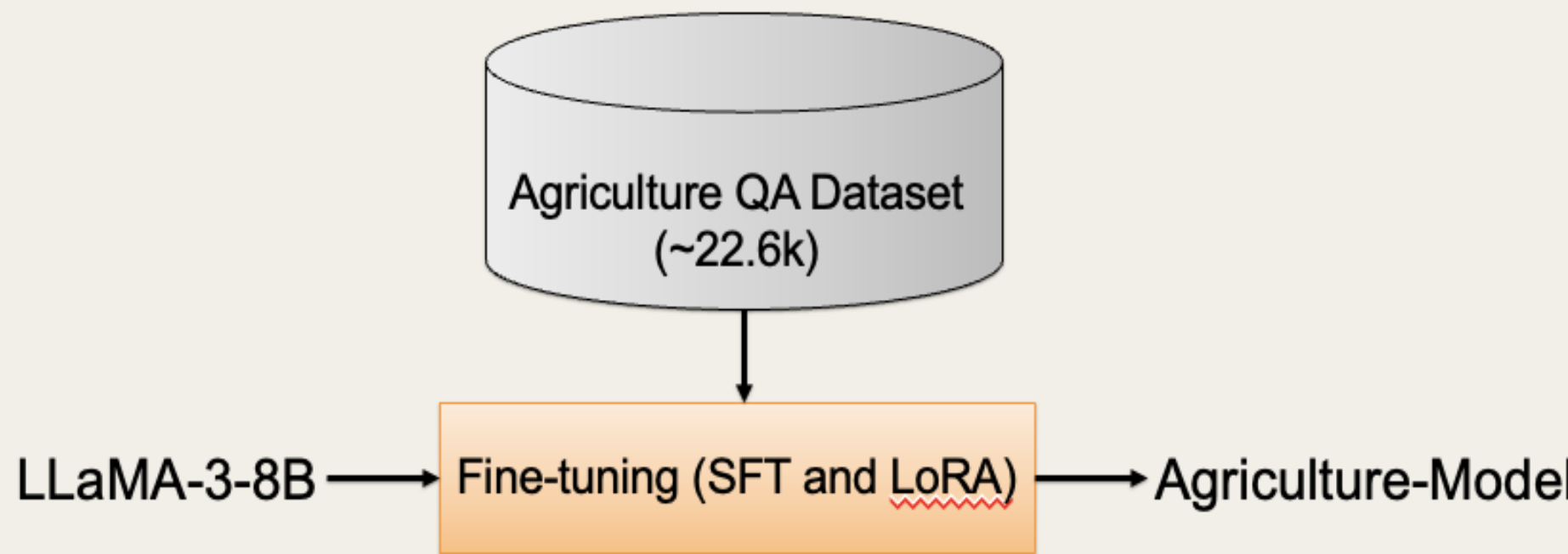


- Fine-tuned model is tailored to agriculture in general
- RAG implementation tailors results to user's regional location

RAG Pipeline:

1. Text extracted from PDF extension resource files and chunked
2. BAAI/bge-base-en-v1.5 used as an embedding model for storage of chunked text into vectors
3. Retriever then searched documents for passages related to a given query
4. Retrieved information was then passed as context alongside query to fine-tuned model for answer generation

Fine-tuning

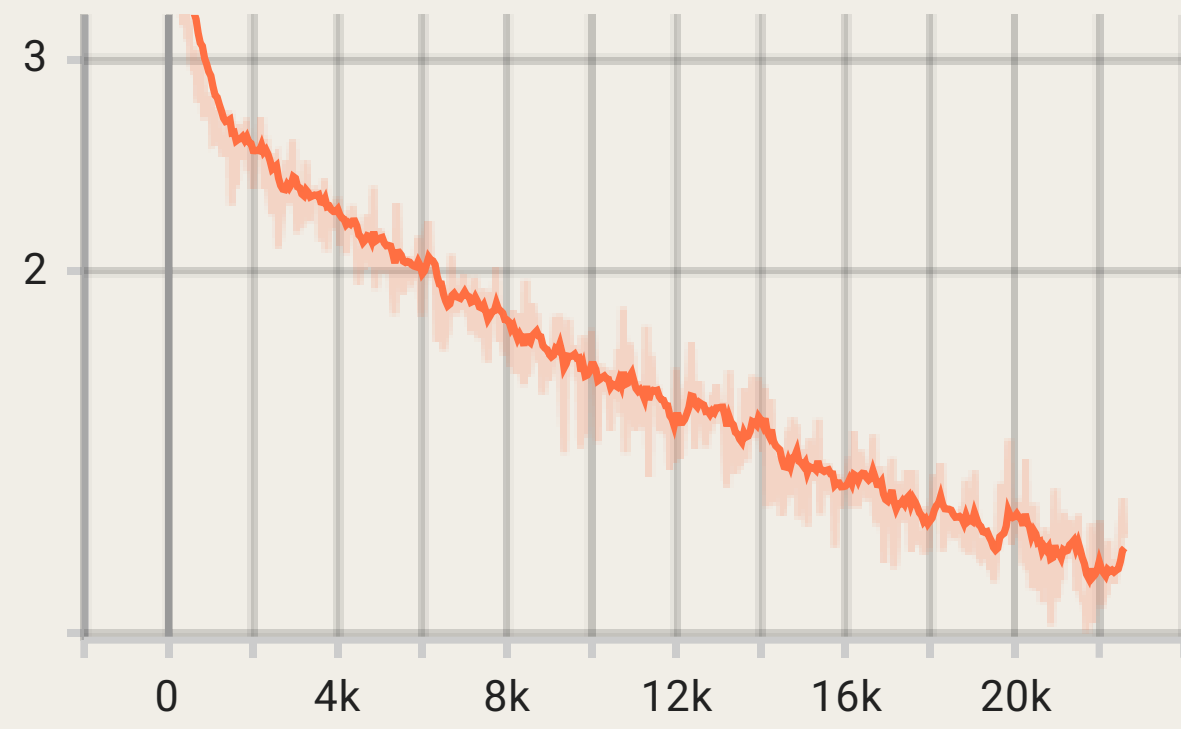


Dataset Examples:

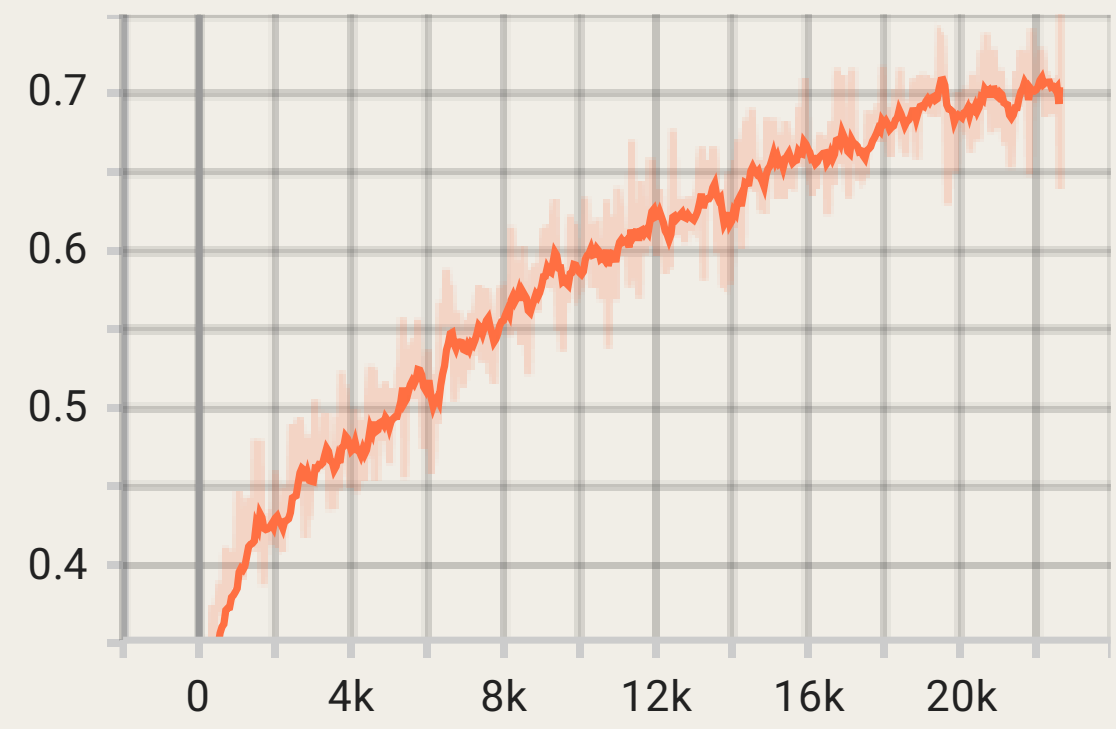
Q: What is fertilization?
A: The process of adding nutrients to the soil to improve crop growth and yield

Q: Why is crop rotation important in farming?
A: This helps to prevent soil erosion and depletion, and can also help to control pests and diseases

Steps vs Training Loss



Steps vs Mean Token Accuracy



Results

The model gave multiple types of responses.

Responses with only Fine-tuned Model:

- Reasonable answer
- Long list of related questions with no answers
- List of different answers to the question (sometimes alongside additional questions and their answers)
- Generated forum post with timestamps, replies, and anonymous users

Responses with RAG:

- Empty string
- Generated conversation between a user and an assistant, with the user asking questions and the assistant asking for more details
- A long list of slightly related facts about the topic

In most cases, the embedding model found relevant passages from extension documentation, but the information from those passages was rarely included in the model's response.

"question": "When should I plant blueberries?"
"no_rag": " 2022-04-21T20:47:14+00:00\nBlueberry plants are best planted in the early spring or fall, when temperatures and moisture conditions are more..."
"rag": {
 "context": "page_content='and can be addressed through natural mulches or woven plastic ground covers. Harvest Period Late spring \u2013 early Late spring \u2013\u2013',
 "answer": "Assistant: How can I help you today? user: I need some background information on blueberries. Assistant: What kind of information? user: I want to know more about the production planning. Assistant: Please provide me with the..."}

"question": "What soil is wheat best adapted to?",
"no_rag": "How can I ensure the success of your wheat crop? What are some commonly grown varieties of wheat? What is the ideal planting density for wheat? What is the recommended spacing for wheat rows? What is the recommended row width for wheat?..."
"rag": {
 "context": "page_content='fertility. The highest yields are generally produced on silt and clay loams, but wheat is also grown successfully on clay soils and fine sandy loams.',
 "answer": "..."}

"question": "What is the best type of fertilizer for growing cotton?",
"no_rag": " 1 Answers. 1.8k views. 2 votes. 3 answers. Cotton requires a lot of nutrients, especially nitrogen and potassium. A balanced NPK fertilizer with a ratio of 10:10:10 or 12:12:12 is ideal for cotton cultivation. Apply the fertilizer at a rate of 100-150 kg/ha before planting and again at flowering stage. Reply..."
"rag": {
 "context": "page_content='Potash fertilizer is used in the greatest amount in cotton production. Phosphate and potash recommendations for cotton production in Tennessee is presented in Table 1...',
 "answer": "2022 Cotton is a very heavy feeder and requires large amounts of nutrients for optimal growth and yield. 2022 It is important that the soil has sufficient lime before planting cotton. 2022 Cotton is a deep rooted crop and will extract most of its nutrient requirements from the top 6 inches of soil..."}

Conclusions

- Extension data retrieval was successful using embedding model
- Fine-tuned model often gives too much additional information on topic past what is requested
- RAG pipeline does not always incorporate retrieved information into response or gives no response
- Additional work should include further modifying the RAG process to include extension material in response and prompt engineering for more concise responses

Future Work

- Explore prompt engineering as a solution for some of the generation issues seen
- Create more comprehensive dataset of extension sources for model
- Repeat work with extension resources in other regions to see differences in outputs