

Product Recommendation RAG Project Assignment

Overview

This assignment involves developing a product recommendation system using Retrieval Augmented Generation (RAG) in a Jupyter Notebook. You will be provided with an XLSX template containing product data. Note that the template includes some extra fields which you may remove if they are not needed. In this updated project structure, you will build a system that not only integrates embeddings, a vector database, and a language model to generate personalized product recommendations, but also handles general skincare queries via external data sources.

Objectives

- Choose an embeddings strategy/model: Evaluate various embedding models (e.g., Sentence Transformers, OpenAI embeddings, etc.) and decide which best suits the product data.
- Choose a vector database: Select from available options (e.g., FAISS, Pinecone, Weaviate) based on performance and ease of integration.
- Choose a language model: Decide on an appropriate LLM (e.g., OpenAI GPT, Hugging Face models, etc.) to interpret queries and generate output.
- Implement hybrid query routing: In addition to product recommendation queries (e.g., a moisturiser under 1200 for oily skin), your system should handle general skincare inquiries (e.g., How can I treat acne) by routing these to a different data source of your choice (e.g., via a web search). You could also prompt engineer to handle alternate queries or add additional datapoints

You may use frameworks like LangChain or LangGraph, or develop your integration independently. You are encouraged to experiment with alternative search methods, multi-LLM/agent approaches, and innovative RAG architectures.

Assignment Tasks

1. Data Preparation:

- Load and preprocess the provided XLSX file containing product information.
- Remove any additional fields from the template that are not required for the recommendation system.

2. System Implementation:

- Design and implement a RAG architecture that integrates your chosen components into a working pipeline for product recommendations.
- Develop logic to detect the intent of user queries. For product recommendation queries, process using the provided product data; for general skincare queries, route the query to a different external data source (e.g., web search) to fetch additional insights.
- Ensure that your Jupyter Notebook is well-organized with clear code and explanatory text.

3. Hybrid Query Routing:

- Implement logic that distinguishes between queries intended for product recommendations and those for general skincare. For example, a query like 'a moisturiser under 1200 for oily skin' should leverage the product XLSX data, while a query like 'How can I treat acne' should be directed to an alternative data source of your choice.
- Document any algorithms, heuristics, or decision rules used to differentiate and route these queries effectively.

4. Demonstration:

Record a screen demonstration of your project. The recording should thoroughly explain your tech stack choices, the design decisions behind your RAG implementation, and your approach to hybrid query routing, highlighting any challenges encountered during development.

Evaluation Criteria

- Clarity and organization of the Jupyter Notebook
- Justification of tech stack and architectural choices
- Effective integration of the selected embeddings, vector database, and language model
- Innovation in experimenting with different search methodologies, multi-LLM setups, and hybrid query routing
- Quality of the screen-recorded demonstration

Submission Guidelines

- Submit your Jupyter Notebook along with the screen recording of your demonstration.
- Ensure that the notebook includes detailed explanations of your approach and the rationale behind your decisions.

Bonus points

- Bonus point if you implement a chatbot with history, i.e. it has context of previous queries (optional), or any other feature you find relevant.

We understand this is a time taking project, we would appreciate if you put in the effort, this can easily be added as a project on your resume regardless.

Good luck, and we look forward to seeing your innovative solutions!