**Nugget Chatbot: Technical Documentation**

**Overview**

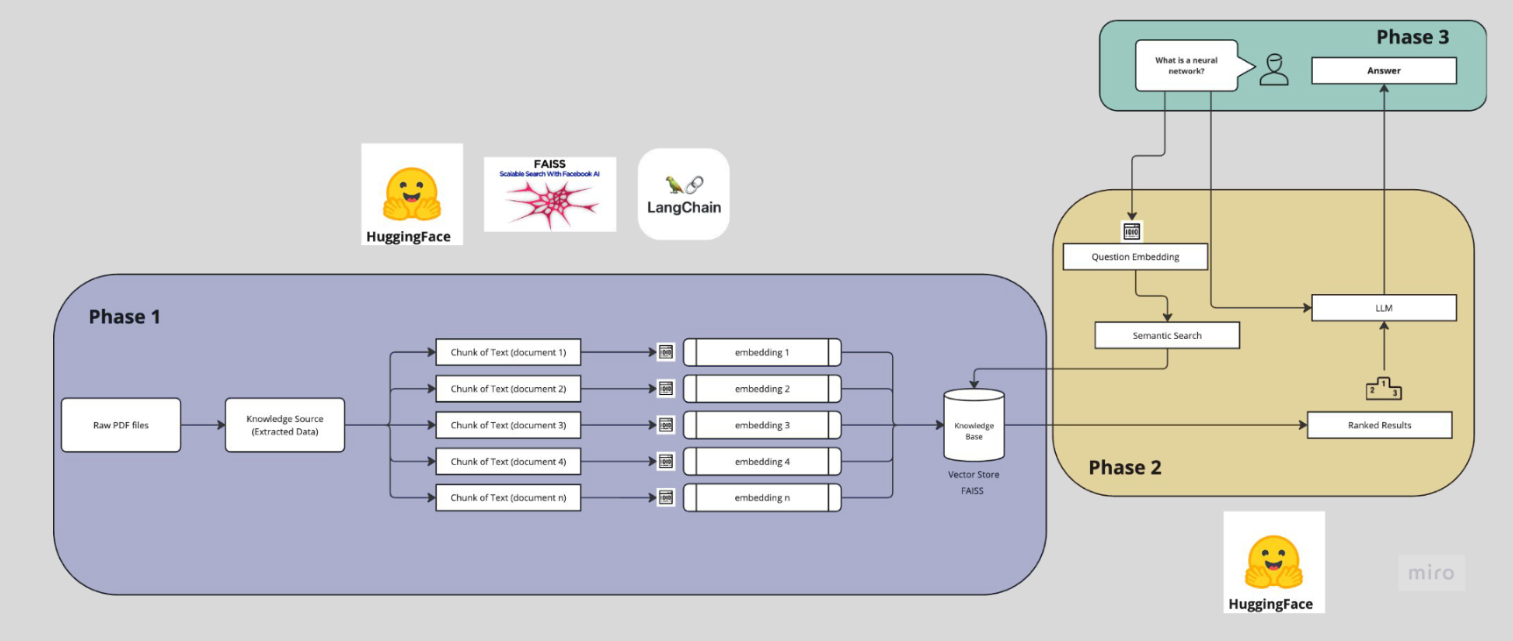
The Nugget Chatbot is a restaurant recommendation system built using Retrieval-Augmented Generation (RAG) to provide intelligent and context-aware suggestions to users based on a custom dataset. The primary source of data is Zomato's restaurant listings, which were scraped, structured into a database, and integrated with a conversational UI to offer an intuitive user experience.

**1. System Architecture**

**1.1 High-Level Components**

1. **Data Scraper**: Scrapes restaurant data from Zomato's website.
2. **Data Processor & Storage**: Cleans, structures, and stores the scraped data in a database.
3. **RAG Model**: Combines a retriever to fetch relevant context and a generator (LLM) to provide natural language responses.
4. **Frontend UI**: A chatbot interface where users can interact with the system.
5. **Backend API**: Serves the frontend, handles queries, and communicates with the RAG model and database.

**1.2 Workflow Diagram**

****

**2. Implementation Details & Design Decisions**

**2.1 Data Collection**

* **Source**: Zomato's public restaurant pages.
* **Tech Stack**: Python, BeautifulSoup, Selenium.
* **Fields Scraped**: Restaurant name, location, cuisine, ratings, pricing, menu items.
* **Data Cleaning**: Normalization, duplicate removal, handling of missing values.

**2.2 Database Design**

* **Type**: PostgreSQL (structured), FAISS (vector DB for semantic retrieval).
* **Tables**:
  + restaurants(id, name, location, cuisine, rating, price\_range)
  + menu\_items(id, restaurant\_id, item\_name, price)

**2.3 RAG Pipeline**

* **Retriever**: Sentence-transformers (SBERT) for dense vector search.
* **Generator**: Mistral model fine-tuned on restaurant query data.
* **Rationale**: RAG allows precise retrieval of relevant info from the database, enhancing factual accuracy.

**2.4 Chatbot UI**

* **Frontend**: Gradio
* **Features**:
  + Query box for user input
  + Display of chatbot responses
  + Fallback suggestions in case of no result

**3. Challenges Faced & Solutions Implemented**

**Challenge 1: Inconsistent HTML Structures in Scraped Pages**

* **Solution**: Implemented fallback selectors and extensive try-catch handling during scraping.

**Challenge 2: Long Response Times for Semantic Search**

* **Solution**: Implemented caching and approximate nearest neighbour search using FAISS.

**Challenge 3: Ambiguous User Queries**

* **Solution**: Incorporated prompt engineering to rephrase queries internally before sending to the retriever.

**4. Future Improvement Opportunities**

* **Multilingual Support**: Enable query handling in multiple languages.
* **Voice Integration**: Integrate with voice assistants for hands-free queries.
* **User Personalization**: Profile-based recommendations (dietary preferences, history).
* **Review Sentiment Analysis**: Add sentiment scores to enhance recommendation accuracy.
* **Real-time Data Sync**: Automatic syncing with updated listings on Zomato.
* **Scalability**: Use of distributed retrieval systems to support large-scale deployment.

**Conclusion**

The Nugget Chatbot leverages the power of RAG to provide personalized restaurant suggestions in a conversational format. With a clean architecture and modular design, it is poised for continuous enhancements, making it a robust and scalable solution for food recommendation use cases.