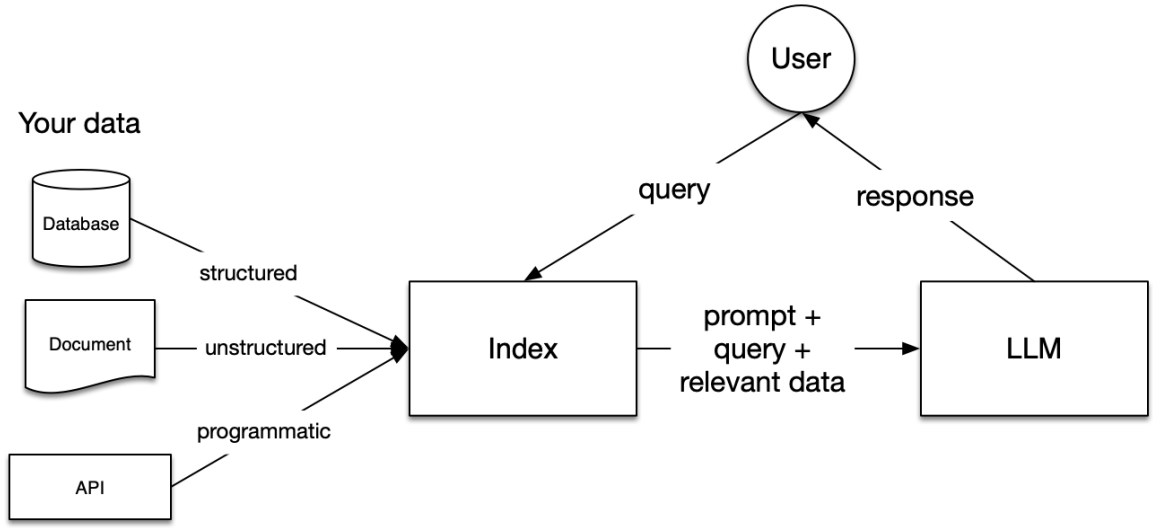
MULTILINGUAL SPEECH RECOGNITION MODEL FOR RAG

RAG



Source :llamaIndex

Procedure :

1) Data Loading
    1. **Dataset**
        (a) The **RAG Mini BioASQ** dataset is a specialized dataset intended for use with retrieval-augmented generation (RAG) models. It comprises passages and questions related to biomedical literature
        (b) Each passage from the dataset is written to the file[txt1.txt]
    2. **Model**
        (a) embeddings and language models using the HuggingFace API, specifically the BAAI/bge-base-en-v1.5 model for embeddings
        (b) HuggingFaceH4/zephyr-7b-alpha model for language tasks,
    3. **Reading dir**
        (a) he code reads and loads data from the directory using the `SimpleDirectoryReader`

2) Indexing and Storing
    1. ChromaDB client to create or access a collection named "quickstart", sets up a vector store
    2. Build a vector index from the loaded data using the specified storage context

3) Querying
    1. Model for Speech Recognation and translation
        a) "openai/whisper-large-v3”-whisper -v3 is sued for Speech recognition and Translation
        b) Audio file as input "/content/tamil_4.weba"- generate result as “study on hidden Markov models”
4) Result
    1. Using query engine (`index.as_query_engine()`) to search based on the text recognized from your ASR pipeline
5) Evaluation
    1. Loading the "rag-datasets/rag-mini-bioasq" dataset, specifically the "question-answer-passages" subset.
    2. Using Sentence Transformers to encode text into embeddings
    3. Computing cosine similarity between the query result (`r`) and the actual answer (`a`) from the dataset.
    4. If the score exceeds a threshold of 0.6, the responses is relevant.
    5. Accuracy Calculation=(`relevant / (relevant + not_relevant)`)
    6. Accuracy=73%