Brance Founding AI Applied Researcher -Hiring Task

Name: Aman Ratnam

Linkedin Profile:

Date Challenge Received: 7 July 2023

Date Solution Delivered: 10 July 2023

1. Problem Statement

The task at hand is to build a chatbot using the Retrieval Augmented Generation (RAG) approach. The chatbot should be able to receive user questions as input and provide answers based on the knowledge provided in a document.

2. Approach

Your approach to the problem. Mention any assumptions made.

1). Document Loading and Splitting - Langchain DirectoryLoader use to load the docs from specified directory , After this document then split into small chunks using **RecursiveCharacterTextSplitter.**

2).OpenAI Embeddings: The code uses the OpenAIEmbeddings module to embed text data using the OpenAI language model. This step is crucial for calculating similarity scores between queries and document passages.

3).Pinecone Vector Indexing: Pinecone is used to create a vector index for efficient similarity search. The code initializes Pinecone with the provided API key and environment. Then, a Pinecone vector store is created using the Langchain Pinecone module, and the document chunks are indexed.

4).Similar Document Retrieval: The code defines a function, get\_similiar\_docs(), which takes a query as input and retrieves similar documents based on the query using the Pinecone vector store. The function returns the most similar documents either with or without similarity scores.

5).Language Model-based Question Answering: The Langchain library is used to load a question-answering chain. The OpenAI language model (GPT-3.5-turbo) is utilized for generating answers based on the retrieved documents and user queries. The get\_answer() function combines the steps of retrieving similar documents and running the question-answering chain to generate an answer to the user's query.

3. Solution

I have found one images to describe the flow of the process



4. Future Scope

yes , we can fine tune the LLM models or we can use HuggingEmbedding for creation of an embedding model.