

Curso de Desarrollo de Asistentes Conversacionales

Proyecto Final

Objetivo

Desarrollar un Asistente Conversacional que permita responder de manera cordial y certera sobre preguntas a una base de conocimientos propia usando (RAG).

Componentes

RAG: Generación aumentada por recuperación.

El usuario podría hacer preguntas sobre información que posiblemente no esté en los datos de entrenamiento del modelo, para ello se crea una base de datos vectorial con la base de conocimiento propia.

Base de Datos Vectorial: Chroma, FAISS, QDrant, Pinecone, etc.

Método de chunking (si fuera necesario)

Modelo de Embeddings y LLM

Framework

Utilizar Langchain y/o Langgraph para el desarrollo de la solución.

Prompt

- Claro
- Simple
- Concreto
- Entrada y Salida

Memoria

El asistente debe contar con memoria, necesario para continuar con la conversación.

Interfase

Se tendría que poder interactuar mediante Streamlit, FastAPI, Botpress o alguna otra herramienta.

Entregables

- Descripción de la solución
- Fuente de Conocimiento con la que se construyó la base de datos vectorial
- Código fuente de la solución, preferentemente en repositorio github.
- Archivo requirements con las librerías necesarias para correr la solución en entorno virtual.
- Se valorará descripción y comentarios en el código.
- Se valorará buenas prácticas de código: clases, funciones, clean architecture.
- Se valorará el uso de tools
- Se valorará que pueda acceder al internet cuando sea necesario
- Se valorará funcionalidades que permitan enriquecer la solución

Fecha de Entrega

El último día de entrega es el 25 de noviembre de 2024.