

Overview del Proyecto

Enfoque

Para desarrollar el chatbot de Promptior, utilicé una arquitectura llamada RAG (Retrieval Augmented Generation), que combina la recuperación de información con la generación de respuestas en lenguaje natural. Usé la librería LangChain, que facilita la integración de estos dos elementos. Cargué la información relevante sobre Promptior desde un PDF, simulando como si fuese una base de datos, y configuré un sistema de búsqueda con FAISS para que el chatbot pueda encontrar las partes más útiles del documento y responder preguntas de manera precisa.

Lógica de Implementación

Recuperación de Información: Cuando el usuario hace una pregunta, el sistema utiliza FAISS para encontrar los fragmentos más relevantes en el PDF de Promptior.

Generación de Respuesta: Luego, el modelo generativo (EleutherAI/gpt-neo-1.3B) transforma ese contexto en una respuesta en lenguaje natural. La pregunta del usuario es procesada, se busca el contexto necesario, y el modelo genera una respuesta específica.

Retos y Soluciones

Algunos de los principales desafíos fueron:

Ajuste de Prompts: Ajusté el prompt (instrucción) varias veces para asegurar que el modelo se mantuviera en el tema.

Manejo de Errores: El modelo no siempre genera respuestas perfectas; a veces la calidad de las respuestas variaba y podía haber errores al procesar las preguntas. Implementé algunos controles para validar y mejorar la precisión de las respuestas.

Limitaciones de Procesamiento: Procesar el texto y manejar la longitud de las respuestas fue otra limitación. Esto me llevó a ajustar algunos parámetros del modelo para mejorar el rendimiento como la "temperature" para regular la creatividad y evitar que demore mucho en responder, al igual que la longitud de tokens, entre otros.

Librerías y importaciones: El uso de librerías tampoco fue tan fácil, decidir qué herramientas usar, hacer bien los imports, usar librerías actualizadas y descargar librerías también fueron un reto.

Enfoque personal

En resumen, este proyecto me permitió experimentar con herramientas avanzadas para la creación de un chatbot capaz de responder preguntas específicas de manera precisa.


Aunque fue un desafío, los ajustes que implementé me ayudaron a lograr un chatbot que utiliza el contenido del PDF de Promtior de forma efectiva para responder a las dudas del usuario. En general fue una buena experiencia ya que fue mi primera vez aplicando RAG y me resultó intuitivo e interesante.

Problemas con Deploy

The screenshot displays the deployment interface for a project named 'ChatBot' with ID '09139ec'. The interface is dark-themed and shows a 'REMOVED' button in the top left. The top right corner indicates the date and time: 'Nov 13, 2024 2:16 PM'. Below the project name, there are tabs for 'Details', 'Build Logs', and 'Deploy Logs'. The 'Details' tab is active, showing a 'Deployment in progress' section with a 'View less' link. The deployment process is listed as follows:

Step	Duration
✓ Initialization	(00:00)
✓ Build	(05:18)
✓ Deploy	(04:41)
✓ Post-deploy	(00:00)

Below the deployment steps, it states 'Deployed via GitHub' and shows a commit message 'add requirements' by 'AleGamme16/ChatBot' on the 'main' branch. The 'Configuration' section at the bottom is divided into two columns: 'Build' and 'Deploy'. The 'Build' column shows the 'Builder' as 'Dockerfile' and the 'Dockerfile path' as an empty field. The 'Deploy' column shows the 'Region' as 'us-west1' and the 'Number of replicas' as an empty field. There are 'Pretty' and 'Code' buttons next to the configuration section.

 ChatBot / 09139ec

REMOVED

Details

Build Logs

Deploy Logs

Filter

Date (GMT-3)	Message
You reached the start of the range → Nov 13, 2024 2:16 PM	
Nov 13 14:26:07	Starting Container
<div>+</div> Nov 13 14:26:45	container event container died
Nov 13 14:26:54	container event container restart
Nov 13 14:26:56	container event container restart
Nov 13 14:27:05	container event container restart
Nov 13 14:27:13	container event container restart
Nov 13 14:27:20	container event container died
Nov 13 14:27:27	container event container died
Nov 13 14:27:33	container event container restart
Nov 13 14:27:40	container event container died
Nov 13 14:27:47	container event container restart
Nov 13 14:27:53	container event container died