

APRENDIZAJE PROFUNDO



<GYM CHATBOT>

Presentan

Erick Machuca Buenrostro (742602)

Ángel Gerardo de la Mora (742927)

Ing. y Ciencia de Datos

Profesor: Iván Reyes

30 de abril de 2024

Informe de Desarrollo del Chatbot de Gym y Salud

Introducción

El objetivo de este informe es detallar el proceso de desarrollo del chatbot relacionado con temas de gimnasio y salud. Se utilizó la interfaz de Gradio para la implementación y se empleó un modelo de lenguaje de OpenAI (LLM) a través de su API. El modelo se alimentó con diferentes textos relacionados a la salud y el fitness para que nuestro chatbot que como se mencionó anteriormente, usando el LLM, poder hacer que sea un experto y un buen consultor relacionado hacia dudas de gimnasio y salud.

Proceso de Desarrollo

1. Procesamiento de Datos

Se utilizó la librería langchain para procesar los datos de texto. Esto incluyó cargar los datos desde archivos en una carpeta (folder_path = "textos"), dividir el texto en fragmentos utilizando RecursiveCharacterTextSplitter, y generar embeddings utilizando OpenAIEmbeddings.

2. Creación del Vector Store

Se creó una base de datos de embeddings utilizando Chroma para almacenar las representaciones vectoriales de los fragmentos de texto procesados.

3. Configuración del Modelo de Chat

Se configuró el modelo de chat utilizando el LLM de OpenAI con temperatura de 0.0 para generar respuestas más precisas.

4. Configuración del Sistema de Memoria

Se implementó un sistema de memoria utilizando **ConversationBufferMemory** para mantener un historial de la conversación y mejorar la coherencia de las respuestas del chatbot.

5. Creación de la Cadena de Recuperación Conversacional

Se creó la carpeta de recuperación conversacional (**ConversationalRetrievalChain**) que combina el LLM, el sistema de memoria y el vector store para generar respuestas coherentes y relevantes.

6. Interfaz de Usuario con Gradio

Finalmente, se implementó una interfaz de usuario utilizando Gradio para permitir la interacción con el chatbot. La interfaz toma texto de entrada y muestra las respuestas generadas por el chatbot. Cuenta con una opción para guardar tanto la pregunta del usuario como la respuesta del chatbot.

Para generar nuestro ChatBot decidimos usar la interfaz de gradio por su uso fácil e interactivo, al igual que recordar el uso gratuito de este.


Decidimos usar el LLM de OpenAI por el hecho de que cuenta ya con una base de datos muy extensa de embeddings para generar la codificación de nuestros textos y con esto poder crear un modelo que aprenda de textos y posteriormente se pueda utilizar como un consultor virtual de dudas del tema de fitness y salud.

Entre los problemas que tuvimos en la elaboración de este fitness chat, estuvo en el momento de tener que pasarle textos para que nuestro modelo se alimentara, ya que queríamos textos que tuvieran sustancia suficiente para que el AI pudiera aprender de ellos. Para esto tuvimos que buscar contenido de varias fuentes y comprender cuáles eran los mejores textos relacionados a la salud.

Posteriormente el problema fue tener que cargar los textos para hacer los splits y obtener los embeddings, debido a que esto lo teníamos que elaborar a mano, pero se creó una función para la solución de esto y fuera más optimo.

En conclusión, la elaboración del ChatBot especializado en fitness y salud, es eficiente en parte gracias al uso del modelo de embeddings de OpenAi, esto ayudo a optimizar el entreno del modelo con la codificación de los textos. EL uso del gpt para crear bots concentrados en un área específica y puedan volverse expertos en esta misma por la alimentación del modelo con textos de esta área y así poder generar asistentes personales para temas cotidianos del día a día.

Referencias

- *Introduction* |  *LangChain*. (s. f.). https://python.langchain.com/docs/get_started/introduction/
- *OpenAI Platform*. (s. f.). <https://platform.openai.com/docs/introduction>

- Rippetoe, M., & Kilgore, L. (2011). *Starting strength: Basic Barbell Training*. Mitchell Beazley.
- Starrett, K., & Cordoza, G. (2015). *Becoming a Supple Leopard 2nd Edition: The Ultimate Guide to Resolving Pain, Preventing Injury, and Optimizing Athletic Performance*. Victory Belt Publishing.
- Williams, M. H. (2002). *Nutrition for Health, Fitness & Sport*. McGraw-Hill Companies.
- World Health Organization: WHO. (2020, 29 abril). *Healthy diet*.
<https://www.who.int/en/news-room/fact-sheets/detail/healthy-diet>
- Stoppani, J. (2023). *Jim Stoppani's Encyclopedia of Muscle & Strength*. Human Kinetics.
- Exton, T., Exton, J., Bridger, M., & Bridger, L. (2017). *Leaner, fitter, stronger: Get the Body You Want with Our Amazing Meals and Smart Workouts*. Random House.
- Fitschen, P. J., & Wilson, C. (2019). *Bodybuilding: The Complete Contest Preparation Handbook*. Human Kinetics.
- Schoenfeld, B. J. (2016). *Science and Development of Muscle Hypertrophy*. Human Kinetics.