Aplicaciones LLM

LangChain: Memory

LLM Apps y Memoria

- La memoria de los Foundation LLMs está limitada por su context window.
- La mejor forma de resolver este problema es utilizando la técnica RAG.
- Además de la técnica RAG, LangChain ofrece varias features para gestionar la memoria de las LLM Apps.

Ejemplos

- Conversation Buffer Memory: almacena los mensajes pasados en el prompt.
- Conversation Buffer Window Memory: almacena los mensajes pasados en el prompt con un límite de mensajes guardados.
- Conversation Token Buffer Memory: almacena los mensajes pasados en el prompt con un límite de tokens guardados.
- Conversation Summary Buffer Memory: almacena resúmenes de los mensajes pasados en el prompt con un límite de tokens guardados.

Otros

 Hay otras funcionalidades avanzadas para gestionar la memoria de Agentes y Entidades que puedes ver en la documentación de LangChain.