

Documento Complejo para Chatbot RAG - IA

Introducción a la Inteligencia Artificial

La inteligencia artificial (IA) es una rama de la informática que busca crear sistemas capaces de realizar tareas que normalmente requieren inteligencia humana. La historia de la IA comienza en la década de 1950, cuando científicos como Alan Turing plantearon la posibilidad de máquinas pensantes. Actualmente, la IA se clasifica en tres tipos principales: IA débil o estrecha, que está diseñada para tareas específicas; IA fuerte, que puede razonar y resolver problemas generales; e IA general, una hipotética forma de inteligencia que iguala o supera la humana en todas las áreas.

Técnicas de Aprendizaje Automático

El aprendizaje automático es un subcampo de la IA que se centra en que las máquinas aprendan de los datos. Existen tres tipos principales:

- Aprendizaje supervisado: la máquina aprende a partir de ejemplos etiquetados.
- Aprendizaje no supervisado: la máquina encuentra patrones en datos sin etiquetas.
- Aprendizaje por refuerzo: la máquina aprende mediante prueba y error, optimizando una recompensa.

Modelos de Lenguaje y RAG

Los modelos de lenguaje son sistemas de IA entrenados para entender y generar texto. Retrieval-Augmented Generation (RAG) combina la capacidad de recuperación de información relevante con la generación de texto, mejorando las respuestas de los chatbots al incorporar contexto específico. Los embeddings son representaciones vectoriales de texto que permiten medir similitudes. Las bases de datos vectoriales, como ChromaDB, almacenan estos embeddings para recuperarlos eficientemente.

Frameworks y Herramientas

- Langchain: framework para construir aplicaciones de IA que combinan LLMs con fuentes externas.
- ChromaDB: base de datos vectorial optimizada para almacenar y buscar embeddings.
- Ollama: plataforma para gestionar y ejecutar modelos LLM localmente.
- Langsmith: herramienta para monitoreo y trazabilidad de interacciones con modelos.

Documento Complejo para Chatbot RAG - IA

Aplicaciones prácticas y desafíos

Los chatbots con RAG se usan en atención al cliente, educación y más. Sin embargo, desafíos como la calidad de datos, costos computacionales y explicabilidad siguen siendo importantes. La integración de múltiples herramientas facilita el desarrollo, pero también requiere un manejo cuidadoso de dependencias y monitoreo constante.