Aplicaciones LLM

App básica: resumir un artículo largo

Problema: el límite de la context window

- Los Foundation LLMs están limitados por su context window.
- La versión gratuita de ChatGPT, por ejemplo, no puede asumir un texto de más de 4097 tokens.
- ¿Qué pasa si queremos resumir un documento más largo de ese límite?

Solución

Utilizaremos una chain predefinida de LangChain para ayudar al Foundation
LLM a resumir un documento que sobrepasa el límite de la context window.

Proceso

- 1. Cargamos el documento.
- 2. Chequeamos su número de tokens.
- 3. Lo separamos en pequeñas partes.
- Utilizamos una chain predefinida de LangChain para enviar las partes a chatGPT y conseguir un resumen del documento.