- The dataset should have at least 30 rows.  
- Make sure your query-response pairs are diverse and are not concentrated on a specific query type, document section or page.  
- Explain why you feel this is a comprehensive dataset to gauge the performance of your chatbot.

Firstly, when i received the take-home project the first idea that I wanted to do is to build the pipeline in such a way that, it can work with any pdf document it gets embedded and gets stored into a vector database and then use the db as the retriever and query it to retrieve documents relevant to the query and then pass it on to the LLM for generation.

Ragas Faithfulness and answer relevancy

A screenshot of a computer

Description automatically generated

{'faithfulness': 0.9246, 'answer\_relevancy': 0.9458}