COREFRAGMENT TECHNOLOGIES

CASE STUDY

RAG Based Chatbot for Document Search

Contact Us

☎ 079 4007 1108

https://corefragment.com

COREFRAGMENT TECHNOLOGIES

CASE STUDY

RAG Based Chatbot for Document Search

Overview

RAG based chatbot is able to answer user queries related to documents uploaded by admin. The client wanted to automate document related queries of users. CoreFragment developed custom chatbot that can support upto 500 documents at a time with 40k tokens storage capacity in memory for context retrival.

Region	Industry
Europe	Ai & ML

Use cases

- Frequent user doubts related to particular document can be resolved by bot
- Need of whole document reading is reduced
- User experience improved when bot clear their doubts irrespective of admin availability
- Information access becomes faster

Contact Us

079 4007 1108

https://corefragment.com

COREFRAGMENT TECHNOLOGIES

CASE STUDY

RAG Based Chatbot for Document Search

Development insights

- Documents are splitted into chunks and their embeddings are stored in vector databases.
- When user enters query into chatbot, it also splitted into chunks. LLMs generate response based on semantic search on query chunks with context retrieving.
- LLM can store the query and response both upto ~40k - 50k tokens. It is used by LLM as context retrieving for further query response.

Technology used

LangChain LlamaIndex Hugging Face aws

pandas Ollama Streamlit

Contact Us

☎ 079 4007 1108

https://corefragment.com