I have done code in google colab   
I need install these packages

Langchain\_community

Pypdf

Faisis-cpu

Langchain–groq  
  
\* I used pyPdfLoader for pdf documentation . Yes i face some difficulties to load bangla language .

\* I used chunking strategy here character based . chunksize = 500 and overlap= 100

\*I used embedings with huggingface because it free . the model capture text with semantic search

\* I used FAISS for store vector . it stores data and give a unique id   
\* Output sometimes can’t get accurately because of language problem   
\* We can improve this performance by using paid version LLM model and better Pdf loader using which support bangla language or we can use multilanguage model