RAG based System approach:
1. Used faiss algo/lib to chunk down the documents to vectorDB
2. Train multinomial naive bayes model to classify the subject of the query to which it belongs to.
3. Check the query whether it is actually a question or something else like command, statement, greetings etc using openai API
4. If query == 'question' and subject == 'Physics' , then fetch context by querying to vectorDB, generate prompt and then pass to the LLM
5. Else: pass the query directly to the LLM
6. Text2voice by gtts(google text to speech) and sarvam api

Additional function that can be added to the RAG based system to make better user experience:
1. Multi-Modal Input: Allow users to interact using various formats, such as voice, text, or images. For example, a user could upload an image and ask related questions about it.
2. Emotional Intelligence: Integrate sentiment analysis to gauge user emotions and adjust responses accordingly, enhancing empathy and user satisfaction.
3. Interactive Learning: Incorporate a feedback loop where the system learns from user corrections or suggestions, refining its responses over time.