**Problem Statement : 5.1 LLM and Gen Ai model like GPT for NLP tasks like Question -answering**

**Title\_1**: Chat with multiple pdf

**Objectives**

* Extract and process text from PDF files
* Split the text into manageable chunks
* Use embeddings and vector store for document retrieval
* Implement a conversational chain to answer user queries

### Technology Stack

* **Programming Language**: Python
* **Libraries and Tools**:
  + LangChain- for text processing tasks such as text splitting, semantic search, and conversational AI
  + PyPDF2- to extract text content from PDF files
  + FAISS- on vectorized text data, enabling semantic search functionality
  + Google Generative AI - for advanced NLP generation tasks,
  + ipywidgets for interactive UI - to create an interactive user

### Data Extraction

utilized to extract relevant information from unstructured data sources such as PDF files

### Text Splitting

* text into chunks using RecursiveCharacterTextSplitter

### Vector Store and Embeddings

* Creating embeddings with Google Generative AI
* Storing text chunks in FAISS vector store

### Conversational Chain

* Setting up a conversational chain using LangChain
* enabling the system to understand user queries, provide contextually relevant responses,

### Question answering

* Input: user need to type the question related to uploaded file
* Output: user get the output relevant to question