

## MINOR PROJECT(18CS64)

College	RV College of Engineering®, Bengaluru				
Department	Computer Science and Engineering				
Course	B.E				
Student Names	1.	Anish Felix Mathias	USN Nos.	1.	1RV20CS022
	2.	Arindham Takur		2.	1RV20CS027
	3	Joel Shaji Mathew		3.	1RV20CS061
Project Title	Interactive Document Summarizer using LLM				

**Signature of Internal Guide**

**Dr. Vishalakshi Prabhu H**

**Assistant Professor**

**R.V College of Engineering**

**Signature of HOD**

**Dr Ramakanth Kumar P**

**Professor & HoD**

**R.V College of Engineering**

# **Project Synopsis**

## **Introduction:**

The problem domain of an interactive document summarizer revolves around the challenge of effectively summarising large amounts of textual information in a concise and informative manner. With the increasing availability of vast amounts of digital documents, such as articles, research papers, and news articles, it becomes difficult for individuals to quickly grasp the main ideas and key points contained within them.

An interactive document summarizer aims to address this problem by leveraging language models and natural language processing techniques. It allows users to input a lengthy document and receive a condensed summary that captures the most important information from the original text. The goal is to save time and provide users with an overview of the document's content, enabling them to make informed decisions about whether to delve deeper into the material or to understand the key points without having to read the entire document. Additionally, the interactive aspect involves allowing users to customise the summarization process, such as specifying the desired length, adjusting the level of detail, or highlighting specific sections of interest.

## **Problem statement:**

This project aims to address the challenge of extracting key insights from a large volume of textual content by developing an interactive document summarizer similar to Blinkist. The system should allow users to upload documents, search for specific books or documents, and engage in chat-like interactions with the content. The goal is to provide users with concise summaries and facilitate efficient access to essential information within diverse textual materials.

## Objectives:

- Enable users to upload documents in various formats, such as PDF, Word, or plain text.
- Develop a search functionality that allows users to search for existing documents by title, author, or keywords.
- Develop algorithms that can process and extract key information from documents to create concise and informative summaries.
- Design a user-friendly interface to display the generated summaries, including the title, author, and main points of the document.
- Develop an interface to communicate with a document as if it was a person.

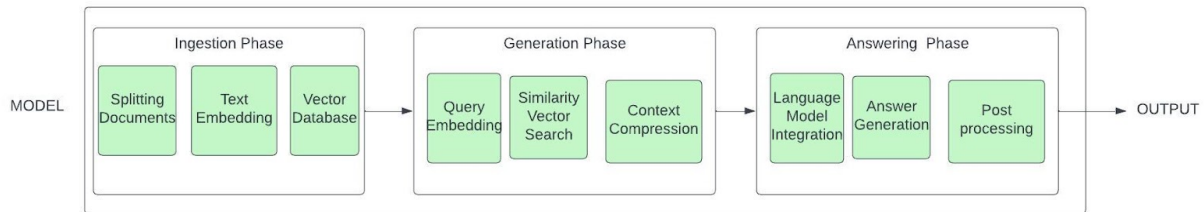
## Methodology:

Our model consists of two distinct phases: ingestion and generation

**Ingestion Phase:** During ingestion, the model receives the entire document collection and employs a semantic and context-preserving splitting mechanism. The resulting text chunks are then transformed into embeddings, which are stored in a vector database for efficient retrieval.

**Generation Phase:** In the generation phase, the user's query is embedded using a similar technique. The model performs a similarity vector search to fetch the relevant text chunks from the stored document embeddings. A context compressor is applied to filter out any irrelevant information, ensuring a more focused input for subsequent processing.

The distilled information, obtained through the ingestion and generation phases, is combined with the user query and fed into our Language Model (LLM). The LLM leverages this refined input to generate accurate and contextually appropriate answers.



## System Requirements:

### 1. Hardware Requirements

- Processor : intel core i5 ,macOS,Linux
- Clock speed : 2.5GHz
- Monitor : 1080\*720 px
- Keyboard : Qwerty Keypad
- RAM : 16GB
- ROM : 8GB
- GPU : Not necessary

### 2. Software Requirements

- Operating System : Windows 10 and above
- IDE : Google Colab Pro
- Front end : OobaBooga
- Python : 3.8 and above
- Framework : LangChain
- Fine tuning : QLoRa

**Expected outcome:**

The expected outcome of our project is to develop a tool that can efficiently summarise books and documents, providing users with condensed versions of the content. By offering a catalogue of pre-summarized books and allowing users to submit their own texts for summarising, the AI aims to save time and enhance information accessibility. The ability to ask questions about the documents further improves user interaction, enabling specific information retrieval. The project aims to create a valuable resource for quickly understanding and extracting key insights from a variety of texts.

**Innovation:**

We're using a unique combination of tools and technologies that no other product has used yet. We predict that because of this we'll be able to match the accuracy of bigger models while having low resource consumption to run on consumer hardware. By not resorting to bigger models, it's possible to run the model locally, without using the internet. This provides a privacy advantage as even confidential documents can be analysed with this tool.