**PDF SUMMARIZATION FOR EACH CHAPTER:**

1. First the first 5 or 6 pages of the PDF is being fetched from the overall PDF and then converted to string.
2. This string is then fed to the GPT 3.5 model to extract Chapter Wise Page Numbers which is then stored as a Data Frame
3. Next using this Data Frame’s page numbers of individual Chapter contents are being retrieved from PDF as string and stored separately with Chapter in a Data Frame.
4. A Summary Chain using Langchain framework is being set up and also a Text Splitter is setup to split each chapter content into multiple chunks.
5. Each chapter content present in the Data Frame is being sent to the summary chain separated via Text Splitter into multiple chunks and then summarized.
6. This summarized content along with chapter name is saved in a JSON