The FinalReportContent.zip contains all the structured code

Folder Structure:

1. Data_Post-Restructuring_with_Additional_Contextual_Tags_for_Question_Answer_Pairs → Adding Contextual Tags and Data Processing
2. GeminiPrompt_QuestionGeneration_and_Training_Validation_Split → The Gemini Prompt and the Training Validation Split
3. Training and Hyperparameter Tuning → Training and Hyperparameter Tuning

The DeepLearing_FinalProject.zip contains all the **experimental code** that we wrote for our project, additionally, it also contains the ***pdf extracted data* (input to the above folder structure)**

Where to find the pdf extracted data?

Since we wrote scripts on top of MINERU, the extracted data jsons can be found under similar hierarchy:

```
mineru_extract-kit>chapter1-6>MIT6390Notes_Chapter{i}>auto>MIT6390Not
es_Chapter{i}_content_list.json
```