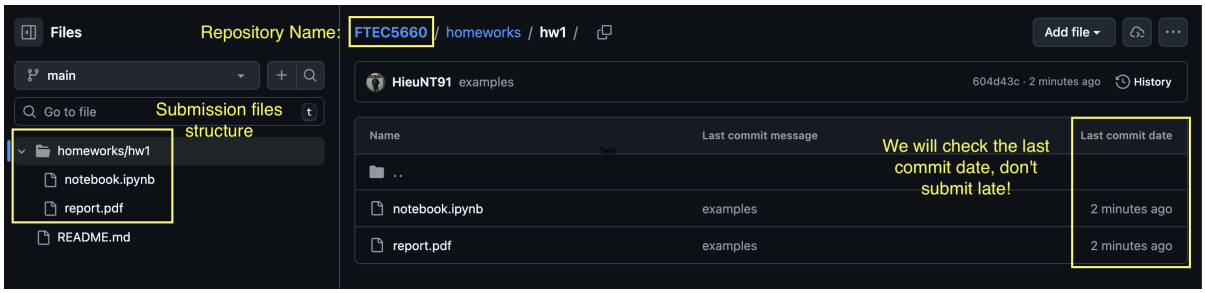
Name: Student ID:

Agentic AI for Business and FinTech (SEEM5660)
Individual Homework 01, Due date: 30 January, midnight.

Instructions

- **Modality:** This is an individual homework. Students are encouraged to discuss with other students, but each student should submit their own work.
- **Input Data:** We posted test image samples in the google drive link on Blackboard. Download the images and upload to your Google Colab session.
- **Resources:** As this assignment utilizes multimodal LLM capabilities, we have provided `Notebook 1` at this link `https://colab.research.google.com/drive/1EMg1-uQwEi8Slnyc3XHz-2UvH7m_KUAo?usp=sharing`, which contains essential starter code. Refer to the sections titled
  1. `Helper functions`
  2. `Image input to Gemini`

  for some basic commands.
- **Submission Requirements:** All students are required to create a public GitHub repository for this course and upload their submission files to the GitHub folder. The GitHub folder should contain at least the following:
  - A report file `report.pdf`. The report file should be maximally 5 pages, documenting your solutions and
  - The python implementation of your solution. For example, you can submit a `.ipynb` file. To download ipynb file, in Google Colab, choose File → Download → Download .ipynb



For students unfamiliar with GitHub, we recommend the following introductory tutorial: `https://www.youtube.com/watch?v=-RZ03WHqkaY`.

1 Practical Problem

Your product (model) should be able to do the following:

- It takes as input *several* images of the supermarket bills, plus one user query (in the form of text).
- There are two possible queries that the users are interested to know:
  1. Query 1: How much money did I spend in total for these bills?
  2. Query 2: How much would I have had to pay without the discount?
- The model should have the capacity to reject irrelevant queries.

2 Evaluation

At test time, we will input new images of the supermarket bills, and ask 10 random queries (covering Query 1, Query 2, and out-of-domain queries). We record how many times your model answer correctly. That gives you the final score for this homework.