Weekly Report

3/5

1. Summary of Weekly Progress

We discussed the architecture and tech stack: the frontend uses React, TypeScript, and Tailwind CSS, while the backend uses FastAPI. For the LLM, we chose Florence-2-large for image-to-text and Llama-3.3-70B for text-to-text. We began with the DDPM paper, as it's key to understanding Stable Diffusion

2. Task Assignments and Contributions

Team Member	Assigned Tasks	Completion Status
徐浩哲(01157030)	1. learn to use react and make a simple web page 2. choice model meeting(jointly) 3. learn basic about uvicorn ASGI server	Complete
翁子翔(01157048)	1. Try to create a simple frontend prototype 2. Read *Denoising Diffusion Probabilistic Models* (Pt.1, Pt.2) 3. Read *Denoising Diffusion Probabilistic Models* (Pt.3 to 6) 4. Test Florence-2-large with Hugging Face	1. In Progress 2. Complete 3. Pending 4. Complete
蔡豐蔚(01157010)	1. Learn how to use Git command upload projects on GitHub 2. Learn how to use FastApi	1. Complete 2. In Progress

2.1. Comments

Comments
use tecticks learn on web to write a easy product of todo list using React. for future backend developing, learn how to use uvicorn ASGI server with python fast api
I try to build a small frontend webpage that allows users to input two images. I attempted to read the *Denoising Diffusion Probabilistic Models* (DDPM) paper, which contains many mathematical formulas and underlying logic that require a long time to understand. So, I started with Part 1 and Part 2, taking notes along the way. When I finish reading the entire paper, I will provide a full summary.
We want to use GitHub to manage our projects, first we use Git command to upload projects and then we can use GitHub Desktop, which is more easier to use. I tried to learn how to write backend routes with FastAPI, but it’s a bit difficult because I’m not sure if the route I set is correct.

3. Challenges and Issues Faced

- Frontend and backend connection
- No yet understand the basic of the backend
- Denoising Diffusion Probabilistic Models contain a lot of mathematical formulas that are difficult to understand
- A single model should be able to run, but we’re unsure if the hardware can handle all of them together

4. Next Steps and Goals for Next Week

We plan to integrate the front-end and back-end, including models such as image-to-text and LLMs, into the web application. This will serve as the prototype and testing foundation for our project. We also need to continue reading the paper in order to better understand how diffusion models work.

5. Additional Notes/Comments

We've finalized the project topic. The paper reading is challenging and requires a lot of time to review other tutorials, understand the concepts, and implement them