

Final Project –

Generating Product Image from Customer Reviews

This project invites you to delve into the intersection of text and image generation. The goal is to use a LLM to extract valuable information from the textual content of the product customer reviews, followed by leveraging diffusion models to re-create product images based on the customer reviews. This task not only tests the generative capabilities of AI in its understanding and interpretative abilities in extracting visual cues from texts, but also its ability in producing relevant visual content based on these visual cues.

Instructions:

Q1. Product Selection and Customer Review Data Collection (5').

- Select 3 different products from different categories on a digital marketplace (e.g., Amazon).
- Please take into consideration different factors (e.g., product categories, popularity levels) when making your selection.
- Explain the rationale of your choices.
- Collect the corresponding product descriptions (textual content) and customer reviews (textual content) for each product.

Q2. Analysis of Customer Reviews with LLM (7.5').

- Use an LLM API to conduct text analysis to extract valuable information from the textual data collected above and build a more holistic understanding about the product.
- Some analyses that are relevant to consider include, but not limited to, for example, text summarization, extraction of particular product features (e.g., visual information), topic extraction, or sentiment analysis.
- Conduct your analyses using prompt engineering (with different prompt strategies), RAG, or a combination of both.
- When doing the analysis, you may need to consider different documentation chunking strategies given the input token limit LLM API has.
- You could consider using a vector database to store your text embedding if necessary.
- You also need to think about what is an effective output from this step, given that your goal next step is to send this output into the diffusion model for meaningful product image generation.

Q3. Image Generation with Diffusion Model (7.5').

- For each product, based on the information extracted from the product description and customer reviews, craft prompts to guide the image generation process effectively.
- Use **two different image generation models** (e.g., OpenAI's DALLE, Midjourney, Stable Diffusion) to generate 3~5 images for each product based on your crafted prompts. Experiment with different prompts and settings to best visualize what you believe is a good illustration of the product based on product description and customer reviews.
- If necessary, iterate on your prompts based on initial results to refine the illustrations.
- Compare AI-generated product images with the actual product images posted in the real world. Are they similar or different? In what dimensions? Do you think AI is able to illustrate the products well? Why or why not?
- Do you see any significant differences across different image generation models? Which one do you like better? Why?
- Provide analyses and explanations of your findings.

(Bonus) Q4. Build an AI Agentic Workflow (5').

- Design and build an AI Agentic workflow to connect all the above steps.

Deliverables:

- Submit a final report documenting all your analyses, findings, and explanations.
- Include links or files for any data you collected, code you wrote, prompts you designed, and images generated.
- In your final report and final presentation, include a reflective discussion on the main challenges faced and lessons learned.
- Every team only needs to submit **one final package**.
- Every team will do a **15-20 mins presentation** during week 7.

Deadlines:

- Final presentation: 12/01 and 12/03
- Final report and package due: 12/05.

Grading

Grade Component	Weight of Total Grade 20%
Q1	5%
Q2	7.5%
Q3	7.5%

Q4

5%

Grade Rubrics	Share of Assignment
Have you understood the question correctly?	10%
Experiment Design (e.g., <i>Creativity and diversity in product selection, prompt design, novelty in tasks</i>)	20%
Analytics (e.g., <i>Is there a clear rationale for why certain things worked or didn't work?</i>)	30%
Insights (<i>Do you arrive at meaningful and informative insights?</i>)	30%
Scientific Rigor (e.g., <i>Code quality, documentation, and reproducibility of the results</i>)	10%