Multimodal Image Generation: Reflection & Practice Worksheet

Learning Objectives

- Understand multimodal learning using text and image inputs.

- Generate images using Stable Diffusion in Python.

- Reflect on how model parameters and prompts influence generated outputs.

Reflection Questions

1. What is multimodal learning, and how is it demonstrated in this project?

2. What role does the text encoder play in generating an image?

3. How does changing the prompt influence the final image?

4. What does the 'guidance_scale' parameter control?

5. What ethical implications arise when generating synthetic media?

Exploration Prompts

Try editing the notebook prompt with creative examples:

- A dragon reading a newspaper

- A robot teacher in a futuristic classroom

- An underwater city with glowing coral buildings

What changes do you observe in structure, lighting, or theme?

Optional Challenge

Modify the code to compare two prompts with and without guidance_scale adjustments. Record how closely

each image aligns with your expectations and explain why you think that is.