TEXT TO IMAGE

A PROJECT LEVERAGING DIFFUSION MODELS FOR AI GENERATED IMAGES

PROBLEM STATEMENT

What is Text-to-Image Generation?

- Converting natural language descriptions into images using AI.
- Powered by diffusion models (like Stable Diffusion).

Objective of the Project

- Implement a Stable Diffusion pipeline using Hugging Face's diffusers library.
- Experiment with different parameters to improve image quality.
- Helps in art creation, story-telling, design automation and AI research.

APPROACH

Model Selection

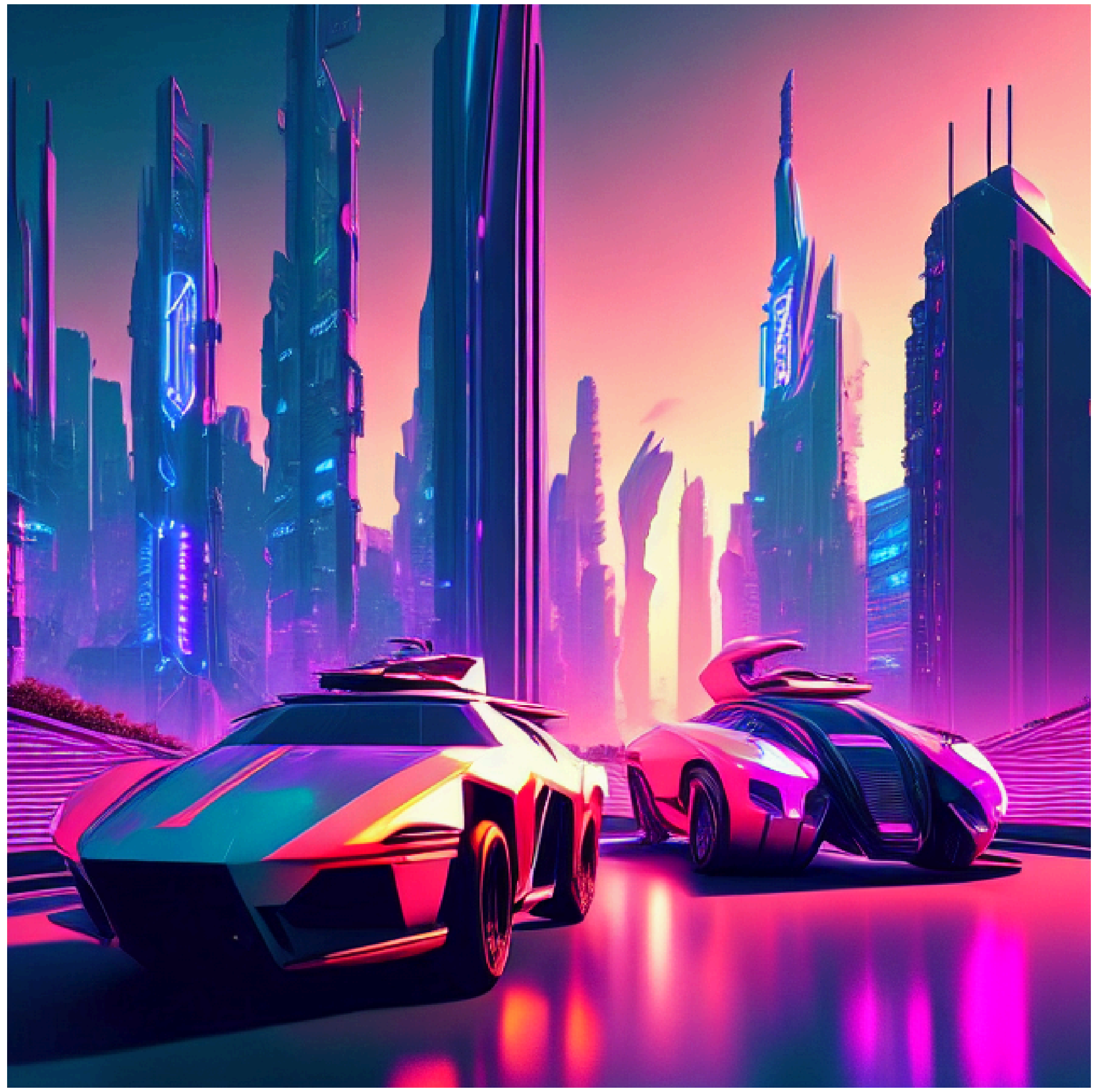
Used stable Diffusion models like :

- dreamlike-art/dreamlike-diffusion-1.0
- stabilityai/stable-diffusion-xl-base-1.0

Image Generation pipeline

Defined a function generate_images() to:

- Accept a text prompt.
- Adjust generation settings (like num_inference_steps, height, width).
- Use Matplotlib to display results.



A futuristic city skyline at sunset, flying cars, neon signs, cyberpunk aesthetic, ultra-detailed, cinematic lighting.

Hyperparameter tuning

- Inference steps (num_inference_steps → affects detail).
- Image size (height, width → controls aspect ratio).
- Multiple images (num_images_per_prompt → generates variations).
- Negative prompts (negative_prompt → removes unwanted elements).

Results

Image Quality Improvement with different parameters.

- More inference steps = Better quality but longer processing time.
- Adjusting image size = Controls resolution.
- Using negative prompts = Reduces unwanted artifacts.
- Future enhancements includes addition of UI for user input.