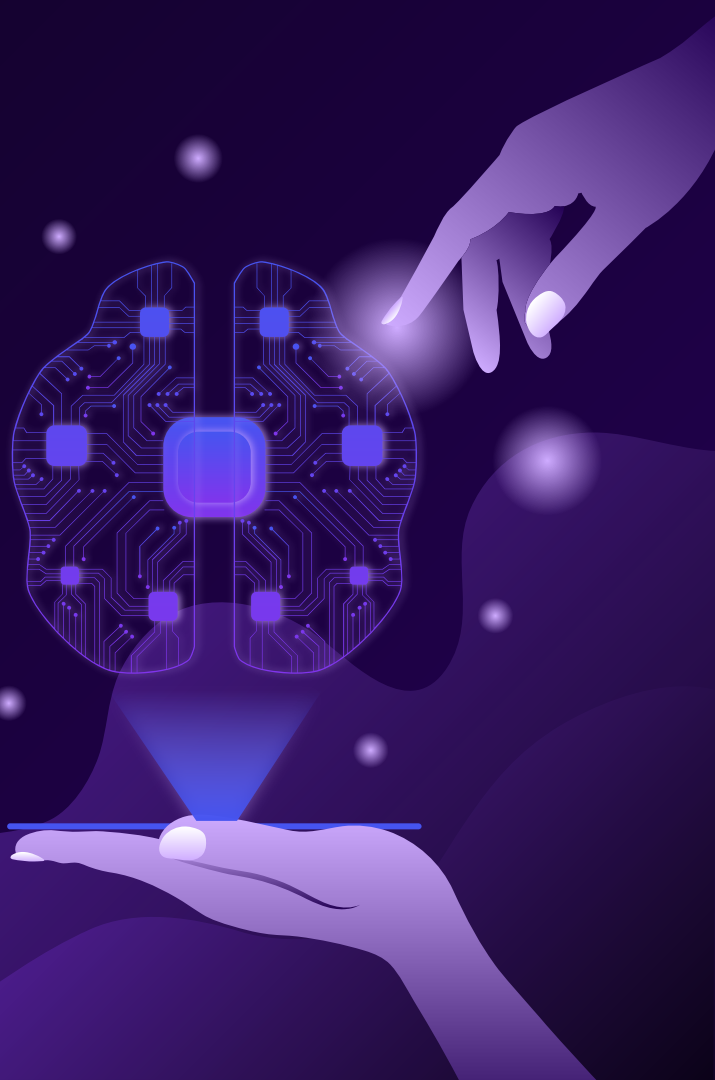
Computer Vision: Text-to-Image Generation

A Comparative Analysis

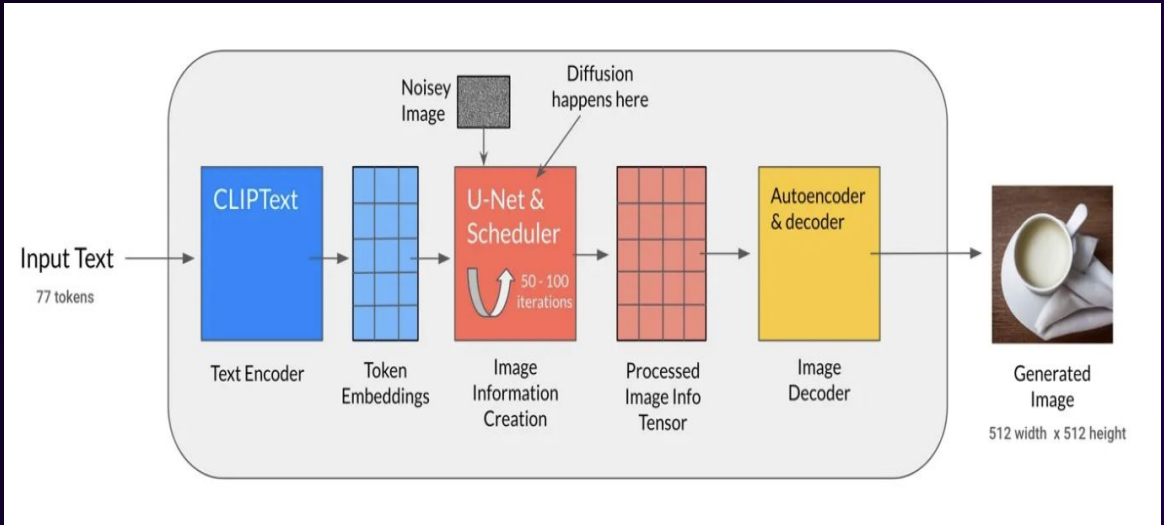
By Big Data Bandits -
Shriya Yegalapati,
Blessy Chinthapalli,
Rachita Harit,
Meghana Kanthadai



Business Objective

- Explore how can we utilize Diffusion in Python to discover and craft stunning images.
- Demonstrate Stable Diffusion 2 advantages over Stable Diffusion 2.1 and DALL-E.
- Sell this as image-on object generation model

Tools and Techniques



Architecture:

1. Autoencoder
2. U-Net
3. Text Encoder

Models Used

1. **Stable Diffusion v2** - Improvements to image quality, conditioning, and generation speed are made.
2. **Stable Diffusion 2.1** - Optimized for speed with AI Template and supports all input shapes up to 1024×1024.

Case Studies and Applications

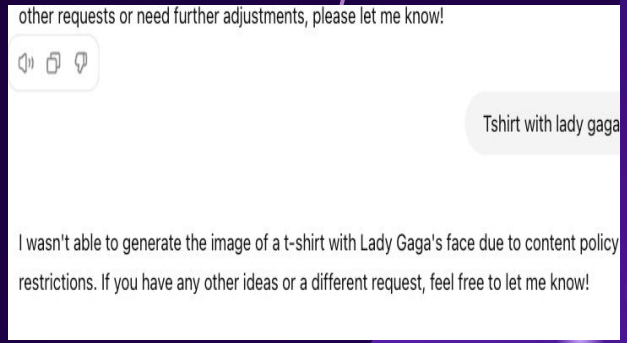
SD 2

Prompt 1: "T-Shirt with Lady Gaga face":

- **DALL-E:** No Design.
- **Stable Diffusion 2:** Realistic, celebrity-like design.
- **Stable Diffusion 2.1:** Bold, graphic design.

SD 2.1

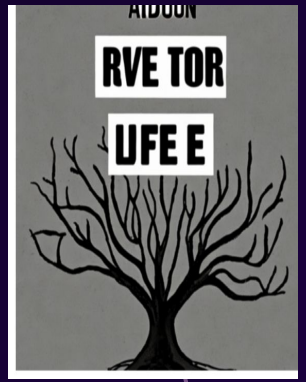
DALL-E



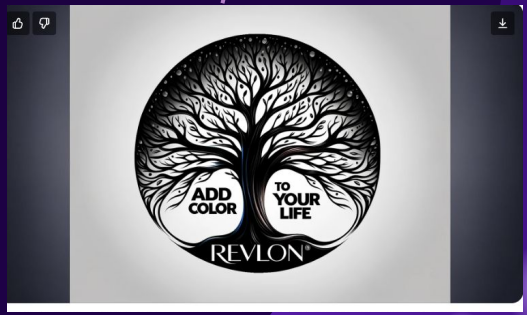
SD 2

Prompt 2: "Black tree with 'Revlon add color to your life'"

- **DALL-E: Clear, visually appealing design.**
- **Stable Diffusion 2 & 2.1: Struggled with text distortion**



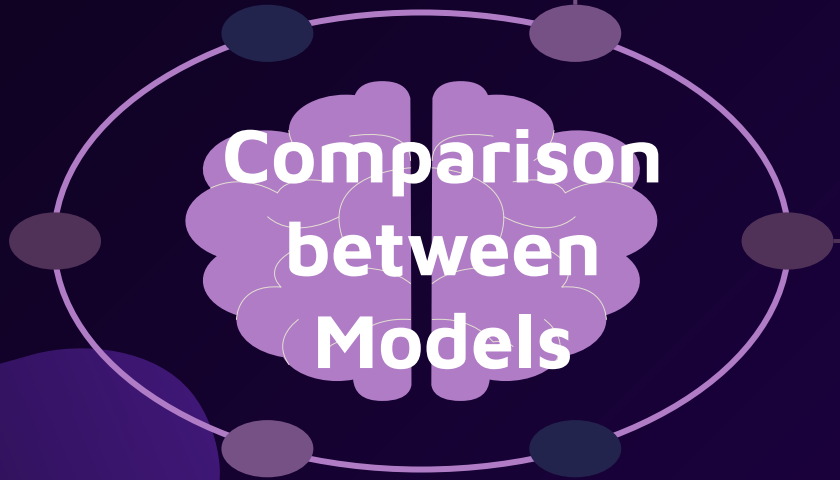
DALL-E



Currently, the Stable Diffusion cannot provide accurate images that have text in them. It gives good results if the prompt is well-written

SD 2.1

Criteria for assessing model performance



DALL·E

- 1280 tokens—256 for the text and 1024 for the image
- Autoregressive
- 64 self-attention layers

Stable Diffusion

- 12 layers
- Diffusion was trained on 2.3 billion images

Fidelity to Concept

Prompt: "phone in water"

DALL-E:	**Stable Diffusion 2:**	**Stable Diffusion 2.1:**
Abstract approach, avoids direct likeness.	Close to celebrity style, realistic within stylized aesthetic.	Similar to SD2 but with a different, bolder style.

Image Quality

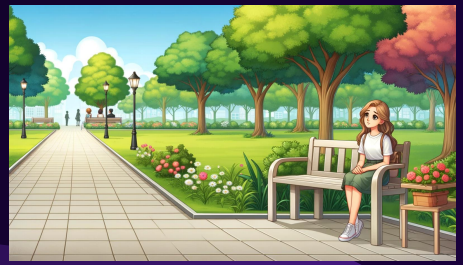
Prompt: "astronaut in space"

DALL-E:	**Stable Diffusion 2:**	**Stable Diffusion 2.1:**
High quality, abstract images.	High quality, clear, and crisp images.	High quality, clear, and crisp images.

Creativity and Uniqueness

Prompt: "girl waiting for her boyfriend in a park"

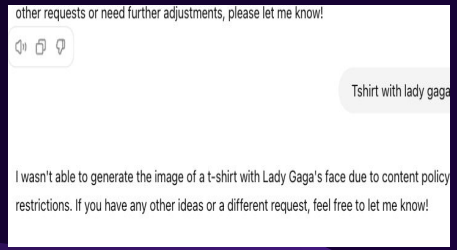
DALL-E:	**Stable Diffusion 2:**	**Stable Diffusion 2.1:**
Unique abstract approach.	Glamorously styled, celebrity-like images.	Bold graphic design with vivid colors.



Realism

Prompt: "Tshirt with lady gaga face"

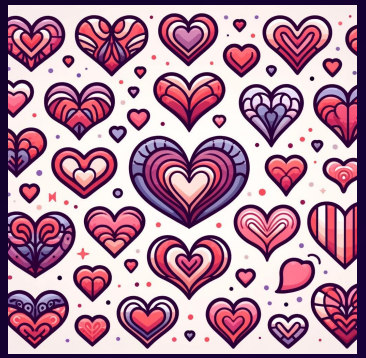
DALL-E:	**Stable Diffusion 2:**	**Stable Diffusion 2.1:**
Not applicable (no direct face).	Realistic within a stylized aesthetic.	Less realistic, more graphic.



Color and Lighting

Prompt: "hearts"

DALL-E:	**Stable Diffusion 2:**	**Stable Diffusion 2.1:**
Subdued abstract colors.	Realistic and naturalistic palette.	Vivid strong contrasts.



Consistency and Professionalism

Prompt: "Painting of kitten in a bowl"

DALL-E:	**Stable Diffusion 2:**	**Stable Diffusion 2.1:**
Consistently abstract.	Detailed, professional quality.	Detailed, strong graphic appeal.

Interpretation of Ambiguity

Prompt: "Woman with her lips closed, stressed out, sitting on a park bench"

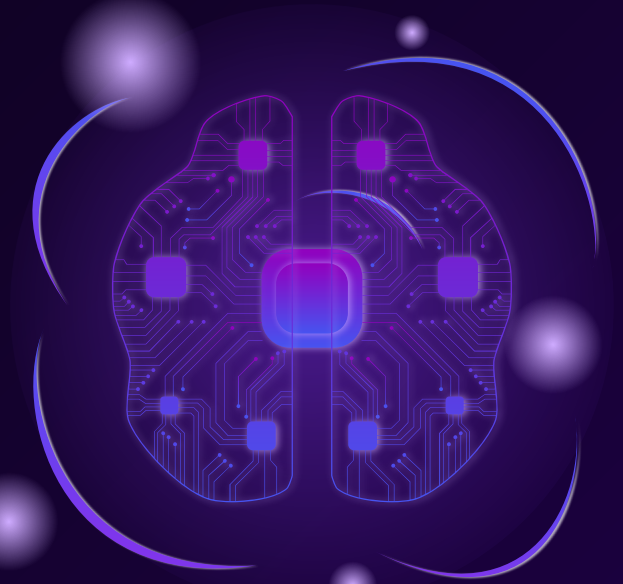
DALL-E:	**Stable Diffusion 2:**	**Stable Diffusion 2.1:**
Cautiously avoids likeness.	Balances creativity with recognizable traits.	More graphic, less focused on likeness.

Marketing Campaign Use Case for Stable Diffusion 2

Use Stable Diffusion 2 to create high-quality, visually appealing images for marketing materials, including social media posts, advertisements, and promotional banners.

Marketing Campaigns

Sample
“beautiful women with cherry red lipstick”



Conclusion: Stable Diffusion 2 stands out as the superior choice for AI image generation due to its balance of quality, creativity, and professionalism.

This model can be used in Marketing Campaigns!

Thanks!