NAME: Leelavathi S

REG NO : 71772117122

COLLEGE: Government College of Technology Coimbatore

Naan mudhalvan -Generative AI Project

PROJECT TITLE

Generate Image uisng Diffusion

AGENDA

Research and Background Understanding

Data Collection and Preparation

Model Selection and Setup

Training and Fine-tuning

Evaluation Metrics and Validation

Results Analysis

Optimization and Improvement:

PROBLEM STATEMENT

- In the era of advanced machine learning techniques, generating images from textual descriptions has emerged as a fascinating and challenging task.
- This project aims to explore the capabilities of stable diffusion pipelines in generating high-quality images based on textual prompts.
- Traditional image generation techniques often lack the ability to produce realistic and diverse images from textual descriptions.
- Explore the impact of various parameters such as prompt length, guidance scale, and model architecture on the quality and relevance of generated images.
- Assess the potential applications of generated images in creative content generation, art, and design.

PROJECT OVERVIEW

- Harnessing stable diffusion pipelines, our project aims to generate high-quality images from textual prompts, advancing the fusion of AI and creativity.
- Through meticulous experimentation, we will refine model parameters to optimize image relevance and diversity.
- Our goal is to pioneer seamless translation of text into captivating visuals, fostering innovation across diverse domains.

WHO ARE THE END USERS?

Content Creators and Designers

Marketers and Advertisers

Educators and Trainers

Assistive Technology Users

Entertainment and Media Industry

E-commerce Platforms

Social Media Platforms

Research and Development

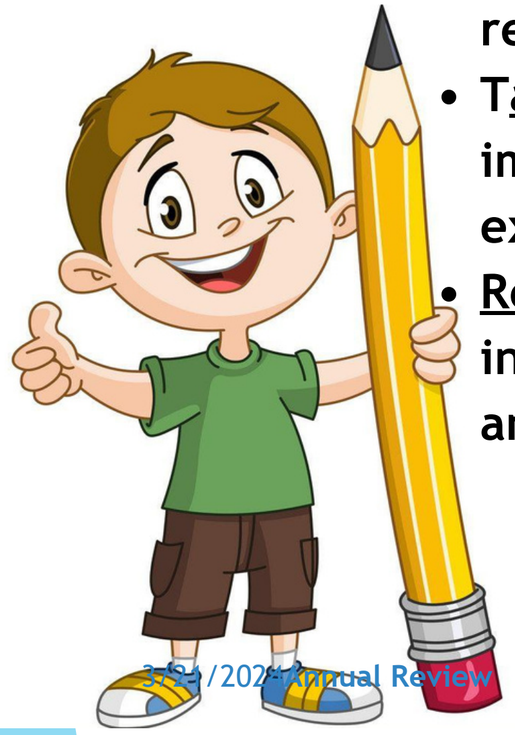
YOUR SOLUTION AND ITS VALUE PROPOSITION

- The proposed solution offers a versatile and intuitive tool for generating realistic images from textual descriptions, catering to a wide range of users across various domains and industries

- By leveraging the capabilities of Stable Diffusion models and NLP techniques, the system enables users to express their ideas and concepts visually, opening up new possibilities for creativity, communication, and innovation.

THE WOW IN YOUR SOLUTION

- Seamless Text-to-Image Conversion: By effortlessly translating textual prompts into lifelike images, our project revolutionizes creative expression and content generation.
- Interdisciplinary Collaboration: Through collaborative efforts across diverse fields, we harness collective expertise to drive meaningful advancements in AI research.
- Tailored Creative Exploration: Empowering users to personalize generated images through simple parameter adjustments fosters limitless creative exploration and expression.
- Real-world Impact and Adaptability: Demonstrating practical applications across industries like art, design, and storytelling, our project showcases its versatility and potential for transformative impact.



MODELLING

Stable Diffusion Model Selection:

- Stable Diffusion models are a class of generative models that leverage diffusion processes to generate high-quality images. These models have been shown to produce realistic images across a variety of domains.
- The selection of a Stable Diffusion model involves considering factors such as model architecture, computational efficiency, and performance on image generation tasks.
- Popular Stable Diffusion models include those based on the Diffusion Probabilistic Models (DPM) framework, such as the models provided by the diffusers library.

Natural Language Processing (NLP) Model Selection:

- NLP models are used to process textual descriptions and extract meaningful representations that can guide the image generation process.

- Commonly used NLP models include pre-trained transformer-based models such as GPT-3, BERT, or T5. These models are fine-tuned on large text corpora and excel at tasks like text generation and understanding.

Training Strategy:

- The training strategy for the combined model involves optimizing the parameters of both the Stable Diffusion model and the NLP model simultaneously.
- Training may involve techniques such as adversarial training, maximum likelihood estimation, or reinforcement learning, depending on the specific architecture and objectives of the model.
- Fine-tuning strategies may be employed to adapt pre-trained models to the task of generating images from textual descriptions.

Evaluation Metrics:

- Evaluation metrics are used to assess the quality, fidelity, and coherence of the generated images.
- Common evaluation metrics include perceptual similarity metrics (e.g., FID score), semantic similarity metrics (e.g., BLEU score), and human evaluation through user studies or crowdsourcing.

RESULTS

Untitled7.ipynb

File Edit View Insert Runtime Tools Help Save failed

Comment Share

+ Code + Text

T4 RAM Disk

```
generate_image(" cricket bat and ball", image_gen_model)
```

13s

100% 35/35 [00:11<00:00, 3.03it/s]

Your paragraph text

RESULTS

Untitled7.ipynb

File Edit View Insert Runtime Tools Help Save failed

Comment Share

+ Code + Text

T4 RAM Disk

```
generate_image(" sun in the blue sky", image_gen_model)
```

16s

100% 35/35 [00:12<00:00, 2.73it/s]

Your paragraph text

DEMO LINK:

https://colab.research.google.com/drive/1_GKyfx6rYqz-imCaqUBAwVsMnZm795hl?usp=sharing