

1. Model Selection/Development:

- I used Stable Diffusion XL (SDXL) for text-to-image generation due to its strong evaluation scores compared to other Stable Diffusion models. It was feasible to use SDXL for both inference and training within a Kaggle environment, where larger models would typically consume excessive RAM.
- I fine-tuned the SDXL(with LORA) model on a sample set of images to generate high-quality, contextually relevant outputs.

Example:

- **Original Prompt:** "MS dhoni running a marathon"
 - **Standard SDXL Output:** Image of MS dhoni running a marathon.
 - **SDXL+Refiner Output:** Image with additional refinement for better visual quality and accuracy.



Fine-tuned Output:

- **Prompt:** "MS dhoni running a marathon"
 - The fine-tuned version showed a marked improvement in quality, providing more accurate and context-specific results that better reflect the scenario described in the prompt.



-
- The model can be further fine tuned with more and better sample images and more number of iterations.

2. Prototype Development:

- I developed a prototype using Gradio to perform inference with the fine-tuned model. This prototype allows users to input prompts and generate images based on the fine-tuned SDXL model. A demo link was provided for testing purposes.
- https://colab.research.google.com/drive/1GZTIzz5_miQNC1onO98-yp35ilr8Y1UK?usp=sharing

3. Deployment:

- At present, deploying this model for widespread use is not feasible without purchasing dedicated GPU resources, as running the model on cloud or local environments requires significant computational power.

4. Business Use case:

- The generative model has potential applications in marketing and advertising. Brands and companies can leverage this technology to fine-tune images for ad campaigns. Actors, like MS dhoni, can be digitally placed in different settings or simulated to appear in various locations, enabling the creation of diverse and tailored marketing visuals without the need for physical photoshoots.

Deliverables:

- You can find the relevant project and code repository here:
- https://github.com/DarrenDsouza47/text_to_img_aliabhatt