Benefits and Challenges of Generative AI

Rakesh (Teddy) Kumar
rakesh.kumar@sri.com

Vice President, Information and Computing Sciences Division,

Lab. Director, Center for Vision Technologies

SRI International
Princeton, NJ

What are the main opportunities introduced by large pre-trained models and Generative AI?

- State of the art performance for Machine Learning systems in Natural Language Processing.
- Much higher performance than previous models
- Ability to do a diversity of tasks using the same model
- Few shot and Zero shot learning of new tasks
- In context learning
- Ability to provide common sense reasoning for a number of tasks.

Making Chat (ro)Bots (Boston Dynamics)

https://youtu.be/djzOBZUFzTw?feature=shared

SayNav: Grounding Large Language Models for Dynamic Planning to Navigation in New Environments

- Use scene graphs to ground LLMs for navigation tasks in unknown large-scale environments
  - Reduce the learning complexity by using a two-level planning architecture - utilizing LLMs to generate a high-level step-by-step plan which can be executed by a pre-trained low-level planner that maps one step into a sequence of primitive actions.

"SayNav: Grounding Large Language Models for Dynamic Planning to Navigation in New Environments", arXiv:2309.04077.


Goal: find a person hiding in the area

What are the main challenges associated with using large pre-trained models?

- Hallucinations, can you trust the output of the large pre-trained model.
- Results can be biased
- Lack of explainability of answers
- Very expensive to train models. Organization would need to invest on server farms as the big AI companies have done.
- Very expensive to do inference with large pre-trained models.
- Many real world applications need to run on the edge.-
- Very difficult to adapt on the edge.

Generative AI for Images

Angel Prompt for Stable Diffusion

Pos Prompts:

full body, walking pose, slow motion, (robot: 1.2), (insanely detailed: 1.5), (highest quality, concept art, 4k), (analog:1.2), (high sharpness), (detailed pupils:1.1), (painting:1.1), (digital painting:1.1), Masterpiece, best quality, (highly detailed photo:1.1), 8k, photorealistic, realistic, real shadow, 3d, (autumn forest with lake background: 1.2)

Neg Prompts:

canvas frame, (high contrast:1.2), (over saturated:1.2), (glossy:1.1), cartoon, 3d, ((disfigured)), ((bad art)), ((b&w)), blurry, ((bad anatomy)), (((bad proportions))), ((extra limbs)), cloned face, (((disfigured))), extra limbs, (bad anatomy), gross proportions, (malformed limbs), ((missing arms)), ((missing legs)), (((extra arms))), (((extra legs))), mutated hands, (fused fingers), (too many fingers), (((long neck))), Photoshop, video game, ugly, tiling, poorly drawn hands, 3d render

Sampler: DPM++ 2M SDE Karras

65 sampling steps, 5.5 CFG scale

Model: RPG-v4

Angel Daruna

Poppy field in Fall

Subhodev Das

Prompt used with DALL-E 3 which is based on GPT 3.

A vibrant field of poppies under morning autumn sky with sun breaking through clouds and fall foliage in the distance.

What is the problem with this image?