MAYANK KUMAR

+918294466003 • LinkedIn

Email Personal Website

SUMMARY

I am always keen on solving real world problem with Deep learning,GenAI ,my previous experience includes working in NLP,LLM tools like langchain and Llama Index, Multimodal LLMS, Diffusion models and Computer Vision

SKILLS

Languages :- Python,C++,JavaScript,HTML,CSS,TypeScript
Frameworks:- Pytorch,Jax,Tensorflow,Keras,LangChain,LlamaIndex,Streamlit,Flask
Tools:- Git,Docker,AWS,Weights & Biases

WORK EXPERIENCE

Machine Learning Engineer ,Oryx AI(Stanford University Stealth Startup)-Palo Alto,California(Remote)

August 2024 - Present

- Worked on Character LLM,Neural Gaussian Fields,SMERF(Efficient radiance fields)
- Worked on diffusion models tools like ComfyUI,regional prompter
- Fine tuned multiple diffusion models with dreambooth,LoRA
- Reduced concept bleeding on multi character image generation using Flux LoRAs

Research Internship,National University of Singapore(Remote)

August 2024-Present

- Working on developing a multimodal AI-powered mental health support system integrating LLMs with real-time speech and facial expression analysis
- We are using llama,audio transformers,computer vision and other tools.
- Guide :-Dr.Tan wee kek

Research Internship,National University of Singapore,Singapore

June 2024

- Our team developed Ikshana, Automated Speech to Sign language system
- Also worked on aws sagemaker
- Guide :- Dr.Tan wee kek ,Dr.Manoranjan dash

EDUCATION

2023-2027

B.Tech in Computer Science Engineering

Bennett University
Specialization: Artificial Intelligence

ADDITIONAL INFORMATION

- **Technical Skills:** Computer Vision,NLP,Cuda,AWS sagemaker,Vision Language Models,Multimodal RAG ,Next.js, React.js,Neural gaussian fields(3d)
- **Certifications:** AWS training and certification (Nus singapore),Big data Analytics using Deep Learning(Nus computing),ML course(coursera)

PROJECTS

Character LLM (GenAI)

- Created character AI that do conversation like specified character and gives empathic response even better than (character.ai).It is trained for role playing ,also used contextual adaptation prompting.
- Fine tuned llama on Experience Reconstruction(a data generation process that can generates detailed and diverse experience data).
- I also deployed this model with interface having option of fine tuning which triggers fine tuning of llama in the backend based on character name and description.
- Github

Anime image generation model (GenAI)

- Fine tuned Animagine XL 3.0 diffusion model on various anime high resoultion anime characters using dreambooth and LoRA fine tuning method.
- Github

Multicharacter image generation using Flux.1 (GenAI)

- Worked on multi character image generation using Flux LoRAs,used LoRA stacking to prompt them together.Prevented concept bleeding using inpainting approach.
- Also used regional prompter extension like ComfyUI and Automatic1111 to prompt different character together.

Ikshana(Speech to American sign language)

- Developed speech recognition model using Bidirectional LSTM,achieved an impressive word error rate approximately 0.2,created pipeline for text to sign language output videos.
- Also worked with CMU Graphics Lab Motion Capture Database

Weights2Weights on custom characters(GenAI)

- Worked on Interpreting the Weight Space of Customized Diffusion Models
- Got results indicating that the weight space of fine-tuned diffusion models behaves as an interpretable latent space of identities
- Created multiple character together using Textual Inversion, which generates personalized images by inverting given images into textual embeddings
- Github

CrossInitialization for expression(GenAI)

- Worked on crossIntialization,a better alternative of textual inversion ,Improved upon this method to successfully edit an individual's facial expression
- Also created multi character with different facial expression together in a image
- Github