Research Engineer | Vision and Language Model Geek

SKILLS

Languages and Frameworks: Python, SQL, Hugginface, FastAPI, PyTorch, DSPy

Technologies: LLMs, Computer Vision, NLP, RAG, Diffusion Models, RestAPI, Docker, Git, AWS, Linux.

EXPERIENCE

Surgical AI Lab (MGH, Harvard)

Research Engineer

- Developed for Binary and Multiclass Video Object Segmentation(VOS) system on surgical video data, achieving sota results, with mean IOU of 92.26% and 74.43% respectively.
- Incorporated SAM-Med2D and XMem++ to do auto-segmentation of anatomy and tools on the videos reducing annotation time for each case from 4 weeks to 1 week.
- Presented two Proof of Concepts, a Vision Language system to generate post-operative reports and summaries using surgical videos and a RAG system to query based on past reports.

Infinstor

Data Scientist

Remote

June 2021 - July 2022

Hyderabad, India July 2019 - April 2021

- Architected end to end Document Understanding pipeline to extract and analyse information from document images using LayoutLM series.
- Deployed an OCR model developed with CNN-LSTM-CTC Loss, to reduce the manual process turn around time by 90%.
- SQL Queries and Cloudwatch logs Anomaly detection using LogBERT. Blog post

Infor

Dev Bussines Analyst

- Led the functional integration of Infor ERPs with Microsoft, and Salesforce CRMs.
- Automated Customer Data Migration, reducing migration duration from 2 days to minutes.
- Trained new developers on Python, with a satisfaction score of 97%

PUBLICATIONS

- Thoracic Surgery Video Analysis for Surgical Phase Recognition. 2nd Workshop on Robot-Assited Medical Imaging, IEEE International Conference on Robotics and Automation, 2024
- Medical Surgery Stream Segmentation to Detect and Track Robotic Tools. *IEEE International* Conference on Artificial Intelligence for Medicine, Health, and Care, 2024
- Utilizing Artificial Intelligence for Surgical Anatomy and Phase Recognition in Thoracic Surgery. IEEE International Conference on Biomedical Engineering Instrumentation, 2024
- Automatic Detection of the Pulmonary Artery During Robotic Right Lower Lobectomy Using Deep Learning. 103rd Annual Meeting of The American Association for Thoracic Surgery, 2023

PROJECTS

nanoChatGPT

Project Link: nanoChatGPT

• Supervised finetuning GPT-2 variants on UltraChat dataset, along with Reward model training on the SFT model with LoRA. Achieved 60-68% accuracy in reward model training on GPT-2 variants. Project Link: DETR

Panoptic Segmentation on Custom Dataset using DETR

• mAP0.5:0.95 of 61% for bbox and Panoptic Quality(PQ) of 54.6% over both things and stuff

EDUCATION

Plaksha University

Technology Leaders Program Awards: Best student committee and Spirit of Plaksha **JNTUH College of Engineering** Bachelors in Technology

August 2022 — June 2023 GPA: 9.53 Credit standing: Gold Medalist 2014 - 2019Percentage: 80

Boston, MA April 2023 - Present