Arjun

551-233-3480 ax2119@nyu.edu Linkedin Github Google Scholar Portfolio

Education

New York University, Courant Institute of Mathematical Sciences **September 2023 – May 2025**
Masters of Science in Computer Science ***3.81/4 GPA*** *New York, United States*

Indian Institute of Technology (IIT) Palakkad **September 2018 – May 2022**
Bachelor of Technology in Electrical Engineering ***8.06/10 GPA*** *India*

Experience

Princeton NLP Group **May 2024 – Present**
Machine Learning Researcher *Princeton, NJ, USA*

- Developing a **multi-agent LLM system** with **Directed Acyclic Graph-modeled** collaborative workflows being executed using **Agent Context Protocols(ACPs)**, dynamically equipping each agent with external tools (web search, Python execution, filesystem, Google APIs, etc) via **Model Context Protocol(MCP) Servers**, enabling adaptive agent interactions and ensuring correct task sequencing, robust error recovery, and structured long-term memory.
- Leading **Reinforcement Learning Fine-Tuning of LLMs (Gemma-3-27B-it, QwQ-32B)** to enhance **agentic deep research capabilities**; developing **verifiable rewards** to **train LLM agents** in effective external tool use (e.g., **search engines**). Leveraging **DeepSpeed** for efficient **scalable distributed training**.

Center for Data Science, NYU **August 2024 – Present**
Machine Learning Researcher *New York, NY, USA*

- Creating **post-training** techniques that improve **spatial understanding** in **DINOv2** and **CLIP/SigLIP** via **language-guided visual grounding**, boosting performance on **dense vision tasks & multimodal** benchmarks.
- Designed a dense dataset pipeline using **LLaVA 3.2 Vision** and **Segment Anything Model (SAM)** to generate local region–caption pairs for **training**; accelerated annotation with **vLLM** for high-throughput processing.

Grossman School of Medicine, NYU **December 2023 – May 2024**
Machine Learning Associate *New York, NY, USA*

- Generated synthetic histopathology images using a **Multi-Modal Model** based on a **Latent Diffusion Model** additionally guided by **Self-Supervised Learning (SSL)** embeddings from **ImageBERT**.
- Achieved ***96.12%*** accuracy in colorectal cancer type classification on real and synthetic images by training a **Vision Transformer**-based encoder with Self-Supervised techniques like SimCLR and StableRep.

Serre Lab **July 2022 – July 2023**
Machine Learning Researcher *Providence, RI, USA*

- Enhanced scale invariant object recognition using biologically-inspired **Convolutional Neural Network**, achieving **30-40%** higher accuracy than **ConvNext** and **ResNet**. Secured **4th** place in the **NeurIPS** Sensorium Challenge.
- Generated synthetic images for color correction tasks, achieving state-of-the-art results on the Color Checker Dataset with a **Recurrent Convolutional Neural Network** model that is **10x** more efficient in size.

Publications

- Arjun et al., **"Agent Context Protocols Enhance Collective Inference"**, Under Review COLM 2025
- Arjun et al., **"HMAX Strikes Back: Self-supervised Learning of Human-Like Scale Invariant Representations"**, Cognitive Computational Neuroscience, 2024 **(Oral Talk)** [Link]
- Arjun et al., **"Subject Independent Emotion Recognition using EEG Signals Employing Attention Driven Neural Networks"**, Elsevier Biomedical Signal Processing and Control (BSPC) [Link]

Projects

Verifiable Rewards: Enforcing Correct Reasoning in LLMs via Reinforcement Learning | *Pytorch, DeepSpeed* **[GitHub]**

- Fine-tuned the **Qwen-7B** LLM with **Group-Relative Policy Optimization (GRPO)**, leveraging **DeepSpeed** for scalable distributed data-parallel training and incorporating **verifiable reward signals** from a larger "teacher" model (**distilled-DeepSeek-R1**). The method boosted **MATH** benchmark performance and demonstrated that enforcing **correct reasoning** measurably improves model **performance** & **generalizability**.

Understanding the Effects Of RLHF and DPO on LLMs | *Huggingface, Python, HPC* **[GitHub]**

- Analyzed the impact of Direct Preference Optimization **(DPO)** and Reinforcement Learning from Human Feedback **(RLHF)** on LLMs' output generalization and diversity by fine-tuning the **Mistral-7B-v0.1** model for summarization tasks using advanced techniques like **PEFT LoRA** adapters and **4-bit quantization**.

Text-To-SQL Context-Aware Query System | *Huggingface, LangChain, ChromaDB, Python* **[GitHub]**

- Created a Text-to-SQL system utilizing LLMs augmented with Retrieval Augmented Generation **(RAG)** to generate context-aware SQL queries. Performed **Parameter Efficient Fine Tuning (PEFT)** of **Llama2-7b** using **LoRA** adapters on WikiSQL & Spider datasets and created a user-friendly interface for querying and interacting with IPEDS.