

# Adithya Iyer

551-344-6506 | [adithyaiyer1999@gmail.com](mailto:adithyaiyer1999@gmail.com) | [LinkedIn](#) | <https://adithyaiyer1999.github.io/>

## EDUCATION

### Courant Institute of Mathematical Sciences - New York University

*Masters in Computational Science*

**Courses Undertaken:** Large Language Models, Machine Learning, Numerical Methods

New York, NY

Aug 2022 - May 2024

### Indian Institute of Technology Bombay

*Dual Degree(Btech+Mtech) in Materials Science*

Mumbai, India

Jul 2016 - May 2021

## PUBLICATIONS

### Cambrian-1: A Fully Open, Vision-Centric Exploration of Multimodal LLMs

Jun 2024

*Principal Investigators : Rob Fergus, Yann LeCun, Saining Xie*

*New York University*

- Led the development of the JAX infrastructure for Vision Language Model **Cambrian-1**, including setting up and optimizing the training runs on TPU-v4 pods.
- Implemented and optimized Fully Sharded Data Parallelism (FSDP) to efficiently distribute the training workload across all TPU machines.
- Implemented Ring Attention and sharded loading/training pipeline to improve scalability and avoid OOMs.
- Designed experiments, conducted comprehensive evaluations on 20+ VLM Benchmarks, and co-wrote the paper.
- Paper titled "Cambrian-1" submitted, accepted by Conference on **Neural Information Processing Systems (Neurips'2025)** as a **Oral Presentation in Vancouver, BC**. Has amassed citations of **270+** as of April'2025

## EXPERIENCE

### Morphic Inc,

*Founding Machine Learning engineer*

June 2024 - Present

*San Jose, CA*

- Responsible for Diffusion Transformer based Video Diffusion Model efforts, including scaling up training infrastructure on multi-GPU nodes, model finetuning efforts, deployment and research efforts related to generative AI.
- Pushed **5+ Video diffusion models** to production on multiple GPUs with **5000+** monthly generations
- Used latest advances in Video Diffusion models to build multi-resolution, multi-time image2Video, video interpolation, extension and sketch guided diffusion models.
- Built evaluation and metrics script for video models, **used widely** across the company for model quality monitoring.
- Optimized production infrastructure for video models to enable **low latency** video generation using model-distillation and compilation

### NYU-x, Computer Vision Lab - Link

*Graduate Researcher, guide: Prof Saining Xie, for reference : [saining.xie@nyu.edu](mailto:saining.xie@nyu.edu)*

Jan 2023 - May'24

*NYC, NY*

- **Recent Work** - creating a from-scratch implementation of multi-modal LLM like Llava in Jax for the team; successfully implemented Fully Sharded Data Parallelism, Ring Attention and Sharded loading/training pipeline.
- Implemented torch-to-Jax model conversion scripts, SPMD mesh selection heuristics based on memory usage and model throughput, and FSDP implementations of various Vision Backbones.
- Past Work - setup **TPU** infrastructure, parallelized training and replicated Diffusion Transformer(**DiT**) based Denoising Diffusion Probabilistic Model on **v4** TPU pods using Jax/Flax framework on **Imagenet** Dataset.
- Envisioned and implemented **TreeDiT**: a **700M Parameter** diffusion timestep based **Mixture of Experts** (MoE) Diffusion Transformer model to speedup training by **20%** and inference by **35%** with minimal impact on FID.
- Replicated Image Generation and Classification via Single Energy-Based (**EGC**) **Model (Link)** in Jax.
- Current research focus involves using diffusion models for **self-supervised** representation learning.

### eBay Inc.

*Applied Research Intern, Machine Learning*

May 2023 - Aug 2023

*Seattle, WA*

- Designed a combination methodology for **unsupervised behavioral embeddings** as features into a pre-existing XgBoost model to predict fraud; achieved a **5 % points** AUC boost by including unsupervised embeddings.
- Finetuned a pretrained **contrastive loss based transformer** model on Account Take-Over(ATO) fraud labels; improved embedding only classifier AUC by **6 % points** across Browser/Mobile Web channels.
- Built **30+ models** across **3 device channels**; identified best model based on precision-recall metrics.
- Performed **swap-sets** analysis for different time periods to quantify benefits of updated model in fraud detection.

## McKinsey & Company, Risk Dynamics Group

Jul 2021 - Mar 2022

Associate Consultant

Delhi, India

- Automated the **credit underwriting process** for the Self-Employed customer segment for a major Indian banking client; created systems to underwrite portfolio worth **60+ mil USD** annually.
- Built **5+** qualitative and quantitative models in **R** from bureau, deposits, bank statements and financial sources.
- Conducted workshops for historical data collection, ensured population stability of variables and executed complete **on-ground implementation** of models; model to be used by **100+** credit underwriters.

## Round Finance

May 2022 - Aug 2022

Backend Development Intern

Mumbai, India

- Designed and implemented backend systems and APIs in **AWS Lambda/MongoDB** by leveraging **Node.js** to enable instantaneous verification of payments in **cryptocurrencies** for the RoundPe payment gateway.
- Ideated, structured and implemented the **crypto-donation link** back-end system to accept crypto donations.
- Setup email, authentication and selective data access to different users/admins by using JSON Web Tokens.
- Conducted load testing of concurrent payments; set up **API rate limits** to protect against malicious activity.

## PROJECTS

---

### Text2KnowledgeGraph - (Website), Github, *website requires some GPT-4 credits* : (Sep 2023 - Dec 2023)

Course: Large Language and Vision Models

NYU Courant

- Built a website which uses **GPT4** API to summarize any input into an interpretable knowledge graph.
- Added capability to summarize inputs such as **Youtube Videos, PDFs, web-pages** and txt files to trees.
- Built a LLM enabled **DFS** algorithm - allowing GPT-4 to traverse graphs and answer questions.

### Structure-Property Relations from Microstructural Images - Thesis, Paper

Jul 2020 - May 2021

Master's Project: Guide: Prof. M P Gururajan and Prof. Hina Gokhale (IIT Bombay)

IIT Bombay

- Awarded **Undergraduate Research Award (URA3)** for excellence in research by the Dean.
- MIST (Microstructure Statistics): A open source library in Python for the analysis of anisotropic microstructures: paper accepted for publication*, library can be found here.

### Budnip : Copernicus Accelerator, European Space Agency (ESA)

Jan 2020 - Jan 2021

Co-founder : Mentored by Dr Alireza Taravat, Deimos Space UK

Remote/Copenhagen

- Won the Oi-X Hackathon** conducted by **Denmark Technical University** and the **Copernicus Program**.
- Built **80+ indices** based on ESA Sentinel-2 satellite spectral images; removed images with excessive cloud cover and performed data augmentation and image-stitching to deal with large raster image sizes.
- Built a **U-Net** based semantic segmentation Deep learning model to classify crops based on vegetative indices.
- Finalist and 2nd runner up** in the **Copernicus Masters - University Challenge 2020**

### Missing Data Importance Weighted Autoencoder (MIWAE)

Sep 2019 - Dec 2019

Guide: Prof. Jes Frellsen

DTU Copenhagen

- Reviewed literature on generative models, variational inference, sampling techniques and data imputation.
- Implemented Importance Weighted Auto-encoder (**IWAE**) by **Burda et al.** in PyTorch.
- Implemented a custom Importance Sampling based lower bound for Missing-at-Random (MAR) pixels in images.
- Built the **MIWAE** model by training on MAR data; produced complete images from incomplete test dataset.

## TECHNICAL SKILLS

---

**Computation related courses:** Optimisation, Data Analysis, Algorithms, Numerical Methods, Machine Learning, Deep Learning, Stochastic Processes, Computer Vision, Large Language Models, Cloud and ML

**Programming Skills/Software Packages:** Python, PyTorch, Jax/Flax, Node.js, React, R, MATLAB, L<sup>A</sup>T<sub>E</sub>X, MongoDB, Git, TypeScript, Django, SQL, TensorFlow, Docker, Kubernetes, GCP