

# PRATEIK SINHA

☎ (+1) 424-535-8133 ✉ [prateiksinha01@gmail.com](mailto:prateiksinha01@gmail.com) 🌐 [prateik-11.github.io](https://github.com/prateik-11) 🔗 [linkedin.com/in/prateik11](https://linkedin.com/in/prateik11)

## EDUCATION

### Carnegie Mellon University

Aug 2025 – Present

M.S. Machine Learning, School of Computer Science, GPA: 4.0/4.0

### University of California, Los Angeles (UCLA)

2020 – 2024

B.S. Double major in Mathematics of Computation & Statistics and Data Science, GPA: 3.901/4.0

## PUBLICATIONS & PATENTS

- *AtmosArena: Benchmarking Foundation Models for Atmospheric Sciences* | Tung Nguyen, **Prateik Sinha**, Advit Deepak, Karen A. McKinnon, Aditya Grover. NeurIPS 2024 Workshops: Foundation Models for Science (FM4Science), Tackling Climate Change with Machine Learning (CCAI). [\[Paper\]](#) [\[Code\]](#)
- *Time-Resolved Neuronal Network Dynamics Distinguish Pathological States in Organoid Models* | Colin M. McCrimmon\*, **Prateik Sinha\***, Qing Cao\*, Tonmoy Monsoor\*, Kartik Sharma, Mehmet Yigit Turali, Ranmal A. Samarasinghe, Vwani Roychowdhury. ICASSP 2026. [\[Paper\]](#)
- *Can Multimodal Large Language Models Understand Pathologic Movements? A Pilot Study On Seizure Semiology* | Lina Zhang, Tonmoy Monsoor, **Prateik Sinha** et al. Submitted to IEEE EMBC 2026. [\[Paper\]](#)
- *Mini-Seizures: Novel Interictal EEG Biomarker Capturing Synchronization Network Dynamics at the Epileptogenic Zone* | T. Monsoor, S. Kanai, A. Daida, N. Kuroda, **Prateik Sinha**, et al. American Epilepsy Society (AES 2024). [\[Paper\]](#)
- *Automated Seizure Classification Using Multimodal Large Language Models* | Lina Zhang, Richard Jiang, Tonmoy Monsoor, Jessica Nichole Pasqua, Colin M. McCrimmon, **Prateik Sinha**, et al. medRxiv preprint. [\[Paper\]](#)
- *Methods And Systems for Automatically Creating Engineering Drawings* | **Sinha, Prateik** & Savant, Shrikant. 2025. US Application No. 63/837,442, filed July 2, 2025. Provisional Patent.

## EXPERIENCE

### Machine Learning Engineer, SolidWorks | Dassault Systèmes

Jul–Sep 2023 (Intern) | Apr 2024–Aug 2025 (Full Time)

- Developed and patented a machine learning method to automatically generate engineering drawings from CAD models. Productionized the method and deployed it on SolidWorks Desktop and xDesign using Python, JavaScript and C++.
- Built ML models to generate and complete simple 3D models, extract specifications of 3D objects from 2D images, and detect features in generic 3D files.

### Data Science Intern | Zelis Healthcare

Jun 2022 – Dec 2022

- Built a pipeline and model to re-price insurance claims based on historical data with Python, R, Snowflake, and Microsoft Azure.
- Developed and maintained dashboards to track insurances claims, their re-pricings, and appeal status. Web-scraped new data using Python (Selenium) and automated cleaning and transforming incoming raw data, as well as updating tables in Snowflake and Azure.

## RESEARCH

### Researcher | Prof. Andrej Risteski, Carnegie Mellon University

Aug 2025 – Present

- Developing machine learning techniques to generate rare transition states in molecular dynamics, reducing reliance on costly simulations.
- Creating synthetic datasets to improve model training and understanding of molecular behavior.

### Researcher | Prof. Aditya Grover, Machine Intelligence Group, UCLA

Jun 2023 – Mar 2024

- Co-authored a NeurIPS 2024 workshop paper (AtmosArena) on benchmarking foundation models for atmospheric sciences.
- Fine-tuned foundation climate models (ClimaX, Stormer) on predicting sparse weather events and subseasonal forecasting.
- Built a framework to assimilate station-level and gridded data to create an improved and more accurate version of the ERA5 dataset.

### Researcher | Prof. Vwani Roychowdhury, The Roychowdhury Group, UCLA

Apr 2023 – Mar 2024

- Developed an approach to identify epileptogenic zones in the brain using network dynamics and machine learning. Co-authored a paper on the same.
- Extended our method to identify genetic mutations in human brain organoids using only 2-photon calcium imaging data. Manuscript in preparation.
- Fine-tuned VLMs to identify epileptic and non-epileptic seizures from videos of patients while providing interpretable, clinical reasoning.

## PROJECTS

### LA Hacks 2023: 3rd place out of 187 teams

April 2023

- Developed an app, people2vec, which matches like-minded people based on their personality by generating an embedding of their YouTube watch history and performing a similarity search against other users. Users can also explore an interactive 3D graph of how their tastes align with their matches.

### HackMIT 2022: 2nd place in ‘Best Use of Blockchain for Social Good’ category

Oct 2022

- Built a crowd-sourced knowledge database, WikiSafe, which stakes all changes/edits to articles on the Ethereum blockchain to create a permanent, immutable record. Also summarizes content, captions images for accessibility, and generates relevant images. Made using React, Flask, Solidity and Web3.js.

## SKILLS

Machine Learning Research/Engineering, AI for Science, Computer Vision (2D/3D), NLP, Data Science, Statistical Analysis, Software Development & Engineering

**Languages & Tools:** Python (PyTorch, TensorFlow, JAX, NumPy), C++, C, JavaScript (React), TypeScript, R, SQL, Snowflake, Azure, AWS, Tableau