# Pranati Modumudi

pranmod01.github.io | (408)-464-8674 | pranati.m@columbia.edu | linkedin.com/in/pranatimodumudi/

## EDUCATION

**Columbia University** Aug 2025 - May 2027 (Expected)

*M.S. in Computer Science, MS Thesis Track in Machine Learning and Computational Neuroscience*

**Relevant Coursework:** Artificial Intelligence, LLM-Based Generative AI, Computation and the Brain, Theoretical Neuroscience

**University of California, Berkeley** Aug 2019 - May 2023

*B.A. Data Science Honors and B.A. Economics | Regents and Chancellor's Scholar | Berkeley SCET*

**Relevant Coursework:** Linear Algebra and Differential Equations; Probability & Probabilistic Modeling; Data Engineering; NLP; Game Theory

## SKILLS

- **ML/DL & RL**: PyTorch, JAX, Scikit-learn, Neural Networks, Reinforcement Learning, Model Training & Evaluation, Distributed Training, MLflow, Weights & Biases
- **LLMs & Generative AI:** Large Language Models, Prompt Engineering, RAG, RLHF, Model Fine-Tuning, MCP, Langchain, HuggingFace, AllenNLP, SpaCy
- **Programming & Systems:** Python, R, SQL, Bash, Java, Git, Docker, Databricks, Snowflake, Big Data Systems, ETL, APIs
- **Neuroscience & Signal Processing:** EEG, Eye-Tracking, Neural Data Analysis, Real-Time Signal Processing, BCIs

## RESEARCH

**Human Decision Modeling | Laboratory for Intelligent Imaging and Neural Computing, Columbia University**

- Developing machine learning pipelines to decode human cognitive states from multi-modal physiological data (EEG, eye-tracking) during decision-making under risk and ambiguity.
- Built preprocessing and feature extraction pipelines for high-dimensional neural time-series data, implementing classification models for physiological markers of risk aversion and uncertainty.

**Pluralistic AI Alignment System | Collaboration with Glen Weyl & Michiel Bakker**

- Developing a prompt-based framework to elicit diverse stakeholder perspectives on contested questions, demonstrating how AI systems can surface value pluralism rather than converging to single consensus answers.

**Undergraduate Senior Honors Thesis | Redwood Center for Theoretical Neuroscience UC Berkeley**

- Developed complex-valued neural network models extending hippocampal place cell representations from rats to bats, investigating cross-species spatial coding differences under Prof. Friedrich Sommer and Dr. Gautam Agarwal.

## INDUSTRY

**Data Scientist I, Disney Streaming,** New York Aug 2023 - May 2025

- Built production ML pipelines deploying Bayesian structural time-series models and gradient boosted decision trees for causal inference on user behavior, serving 150M+ subscribers with real-time prediction latency <100ms.
- Developed a distributed training framework for time-series forecasting models (LSTM, Prophet) processing 10M+ daily events, reducing model training time from 6 hours to 45 minutes via parallelized hyperparameter tuning.
- Engineered reusable Python package for synthetic data generation and continuous treatment effect estimation; deployed via PyPI with automated testing in Jenkins, adopted by five teams for experimental design and model validation.

**Machine Learning Systems Intern, OctoML (acquired by Nvidia), Remote** Oct 2022 - Dec 2022

- Optimized TVM machine learning compiler by running 100+ experiments on CNN models across varying architectures, tuning autoscheduling hyperparameters to reduce convergence time by 10% and improve GPU utilization.
- Diagnosed root causes of performance variance in the TVM auto-tuning pipeline through statistical analysis of compilation traces; findings published in the MLArchSys International Symposium on Computer Architecture 2022.

## PROJECTS

**Multi-Agent SLM Evaluators for AI Slop Detection | AWS Small Language Build Day Hackathon**

- Built an ensemble of four specialized small language models (1-3B parameters) fine-tuned with LoRA on AWS Trainium to detect low-fidelity model outputs, enabling order-of-magnitude cost reduction vs. large-model inference for content quality monitoring.

**AI Capability Terrain Map | Apart Research AI Forecasting Hackathon**

- Developed an interactive 3D early warning system visualizing AI capabilities with real-time Epoch AI benchmark data, logistic growth modeling, and Monte Carlo forecasting across 30+ domains to identify capability trajectories, sinkholes, and emerging risks.