Karthik Prabhu Palimar

Chicago, IL | karthikprabhu22@gmail.com | www.karthikprabhu22.github.io | Karthik Prabhu

Summary

Ph.D. candidate in Physics specializing in statistical inference, machine learning, and large-scale computational workflows. 6+ years of experience in building end-to-end ML systems, from data processing to model development, distributed training, and deployment using Python, PyTorch, and LLM frameworks and Agentic workflows. I enjoy collaborating with cross-functional teams to translate technical requirements into working solutions. I am motivated by challenging problems and I am excited to apply my experience and skills to build solutions that deliver measurable business impact.

Education

Ph.D. Physics *University of California-Davis, Davis, CA* Expected: Dec 2025
B.S. - M.S. Physics *IISER-Pune, Pune, India* 2013-2018

Skills

Machine Learning	Generative AI (VAEs, DDPMs), LLMs, RAG workflows
Frameworks & Tools	PyTorch, Transformers (HF), LangChain, scikit-learn, Pandas, NumPy
Data Engineering	ETL workflows, feature engineering, quality assurance, data modeling
Programming Languages	Python, Julia, MATLAB, SQL

Experience

Doctoral Researcher, *Department of Physics and Astronomy, UC Davis* Sep 2018 - Present

- Developed generative models using Denoising Diffusion Probabilistic Models in PyTorch to simulate complex non-Gaussian foregrounds; preserved target statistics within 25% of sample variance, enabling more reliable downstream analyses
- Released an open-source framework (GitHub) that serves as a $\sim 1000\times$ faster surrogate to N-body runs, making large ensemble studies practical for survey design and bias checks
- Scaled training and inference on Perlmutter (supercomputer) with PyTorch + CUDA + HF Accelerate; converted pipelines to multi-GPU, cutting wall-clock per experiment from over 100s of hours to 30 hours
- Presented results at conferences and journals; maintained reproducible code and documentation for collaborators

Intern, *Handshake AI, Remote* Jun 2025 - Nov 2025

- Designed and evaluated domain-specific prompts to assess large language model (LLM) reasoning in physics
- Analyzed outputs for accuracy, clarity, and depth, informing improvements in scientific reliability of LLMs
- Collaborated with domain experts to integrate findings into design guidelines for newer fellows

Deep Learning Fellow, *Erdos Institute Bootcamp, Remote* Apr 2024 - May 2024

- Built a Retrieval-Augmented Generation (RAG) system that processed 5.5M Reddit posts using GTE-BERT embeddings, LangChain, and sentiment-aware re-ranking; delivered sub-1s query latency
- Improved Mean Reciprocal Rank by 12% across 160 tested configurations through retrieval tuning and evaluation
- Secured first place in the competition for code performance and meeting the KPIs

Selected Publications [Google Scholar]

Learning Correlated Astrophysical Foregrounds with Denoising Diffusion Probabilistic Models
A Generative Model of Galactic Dust Emission Using Variational Autoencoders

Certifications

AWS Educate Machine Learning Foundations
Generative AI with Diffusion Models by NVIDIA
Stanford online Machine Learning course