Prateek Anand

panand2@g.ucla.edu • (408) 431 1325 • https://prateekanand2.github.io/

My research interests are in developing novel AI/ML computational methods. Broadly, I am invested in deep learning, traditional machine learning, and statistical approaches that are scalable and interpretable.

EDUCATION

UCLA | Ph.D. Computer Science **2024 - 2029**

Advisor: Dr. Sriram Sankararaman
I am currently developing improved deep generative models for synthetic data and imputation methods in genetics. I have also recently been exploring masked language modeling for rare feature imputation.

UCLA | B.S. Computer Science **2020 - 2024**

PROFESSIONAL EXPERIENCE

UCLA Computer Science | Sriram Lab **Jun 2022 - Present**

Graduate Student Researcher

- Building deep generative models for genetic variation data (presenting at American Society of Human Genetics in Boston, Oct 2025)
- Developed kernelized, scalable machine learning/statistical software to understand human genetic architecture
- Projects: QuadKAST (co-first author) and FAME

Research Intern (Bruins in Genomics)

- Evaluated novel nonlinear explanation method for machine learning models (form of symbolic regression used to represent complex black box models)
- Presented at UCLA poster symposium

Stanford University School of Medicine | Curtis Lab **Jun 2023 – Jan 2024**

Research Intern (Canary CREST Program for Early Cancer Detection)

- Computational modeling and inference, refined model built using approximate bayesian computation to detect blood cancer early
- Presented at Stanford poster symposium

UCLA Jonsson Comprehensive Cancer Center | Boutros Lab **Jan 2021 - Apr 2022**

Research Assistant

- Updated, maintained, and evaluated CI/CD HATCHet pipeline for use on lab cluster compute server
- Project: Metapipeline-DNA

PUBLICATIONS

QuadKAST | Genome Research **Sep 2024**

- Fu, B.*, Anand, P.*, Anand, A.*, Mefford, J., & Sankararaman, S. (2024). *A Scalable Adaptive Quadratic Kernel Method for Interpretable Epistasis Analysis in Complex Traits*. https://doi.org/10.1101/2024.03.09.584250

Metapipeline-DNA | Preprint

- Patel, Y.*, Zhu, C.*, Yamaguchi, T. N.* et al. Anand, P., ... Boutros, P. C. (2024). *Metapipeline-DNA: A Comprehensive Germline & Somatic Genomics Nextflow Pipeline*. https://doi.org/10.1101/2024.09.04.611267

FAME | Preprint **Sep 2023**

- Fu, B.*, Pazokitoroudi, A.*, Xue, A., Anand, A., Anand, P., Zaitlen, N., & Sankararaman, S. (2023). *A Biobank-Scale Test of Marginal Epistasis Reveals Genome-Wide Signals of Polygenic Epistasis*. https://doi.org/10.1101/2023.09.10.557084

TEACHING

Teaching Assistant **Jan 2025**

- UCLA CM146: Introduction to Machine Learning

Course Reader **Jan 2024**

- UCLA CM146: Introduction to Machine Learning

Course Reader **Jan 2023**

- UCLA CM146: Introduction to Machine Learning

SKILLS

Programming languages: Python, C++

Frameworks: Scikit-learn, PyTorch, Numpy, Pandas, Git

Courses: Machine Learning, Artificial Intelligence, Neural Networks and Deep Learning, Big Data Analytics, Software Engineering and Construction, Data Structures, Algorithms and Complexity, Linear Algebra, Statistics and Probability, Optimization, Machine Learning in Genetics, Algorithms in Bioinformatics

Other: Generative Models, Probabilistic Models, Hypothesis Testing, Cluster Computing

AWARDS

Warren Alpert Computational Biology and AI Fellow: Graduate Training Fellowship 2024

NCI Scholarship: Funding for Canary CREST Research Program 2023

NSF REU Scholarship: Funding for Bruins in Genomics Research Program 2022

UCLA Dean's Honors List: Winter 21, Spring 21, Fall 21, Winter 22, Fall 22, Winter 23

Andy Grove Intel Scholarship: Education Scholarship for Academic Excellence 2020

Valedictorian: Homestead High School 2020