Prateek Anand

Los Angeles, CA | (408) 431 1325 | panand2@g.ucla.edu | 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 **2024 - 2029**

Ph.D., Computer Science

- 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 discrete diffusion and masked language modeling approaches for rare feature imputation.

UCLA **2020 - 2024**

B.S., Computer Science

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 scalable machine learning/statistical software to understand human genetic architecture

Research Intern (Bruins in Genomics Research Program)

- 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 Research 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

PUBLICATIONS

* denotes equal contribution.

- Zhengtong Liu, Boyang Fu, Moonseong Jeong, **Prateek Anand**, Aakarsh Anand, Seon-Kyeong Jang, Aditya Gorla, Jiazheng Zhu, Paivi Pajukanta, Pier Francesco Palamara, Noah Zaitlen, Richard Border, Sriram Sankararaman. Comprehensive gene heritability estimation reveals the genetic architecture of rare coding variants underlying complex traits, *biorXiv*, (2025).
- Boyang Fu*, **Prateek Anand***, Aakarsh Anand*, Joel Mefford, Sriram Sankararaman. A scalable adaptive quadratic kernel method for interpretable epistasis analysis in complex traits, *Genome Research*, (2024).
- Yash Patel*, Chenghao Zhu*, Takafumi N. Yamaguchi*, ..., **Prateek Anand**, ..., Paul C. Boutros. Metapipeline-DNA: A Comprehensive Germline & Somatic Genomics Nextflow Pipeline, *bioRxiv*, (2024).
- Boyang Fu*, Ali Pazokitoroudi*, Zhuozheng Shi, Asha Kar, Albert Xue, Aakarsh Anand, **Prateek Anand**, Zhengtong Liu, Richard Border, Paivi Pajukanta, Noah Zaitlen, Sriram Sankararaman. A biobank-scale test of marginal epistasis reveals genome-wide signals of polygenic epistasis, *to appear in Nature Genetics*, (2023).

PRESENTATIONS

* denotes equal contribution.

- **Prateek Anand**, Anji Liu, Meihua Dang, Boyang Fu, Xinzhu Wei, Guy Van den Broeck*, Sriram Sankararaman*. Deep generative model of genetic variation data improves imputation accuracy in private populations, *American Society of Human Genetics*, (2025).
- **Prateek Anand**, Ryan Schenck, Christina Curtis. Fluctuating methylation clocks and mutational frequencies lead to patient-specific inference of CHIP subclone dynamics, *Canary CREST*, (2023).
- **Prateek Anand***, Aakarsh Anand*, Boyang Fu, Sriram Sankararaman. Explaining potential epistasis in genomic data using symbolic representations of complex black box models, *Bruins in Genomics*, (2022).

TEACHING

- **Introduction to Machine Learning (UCLA CS 146):** Teaching Assistant (Winter 2025)
- **Introduction to Machine Learning (UCLA CS 146):** Reader (Winter 2024)
- **Introduction to Machine Learning (UCLA CS 146):** Reader (Winter 2023)

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)