

NIKHITA DAMARAJU

nikhita@uw.edu • <https://nikhita-damaraju.github.io>

EDUCATION

UNIVERSITY OF WASHINGTON

Doctor of Philosophy (PhD), Public Health Genetics

Courses: Ethical, legal, social implications of Genomics; Principles of Health Metrics, Strategies of Health Promotion, Genetics and Law, Health Economics

Seattle, WA

September, 2022 - present

COLUMBIA UNIVERSITY MAILMAN SCHOOL OF PUBLIC HEALTH

Master of Science (MS), Biostatistics – Statistical Genetics

Relevant courses: Data Science, Machine Learning, Biostatistical Methods-I & II, Statistical Genetic Modelling, Population Genetics, Statistical Inference, Probability, Causal Inference, Probabilistic Graphical Models for Health Data, Epidemiology, Advance Statistical Genomics methods

New York, NY

September, 2020 - May, 2022

INDIAN INSTITUTE OF TECHNOLOGY MADRAS

Bachelor and Master of Science (B.S.-M.S.), Biological Sciences

Chennai, TN, India

August, 2015 – May, 2020

- Minor in Computational Biology
- Government of India DST INSPIRE fellow for five consecutive years given for best academic record

Relevant Courses: Bioinformatics, Biostatistics, Computational Biology, Systems Biology, Quantitative and Population Genetics, Genomics and Proteomics, Immunology, Cancer Biology, Biotechnology in Healthcare

SKILLS AND INTERESTS

Computer Languages

R (rstats, shiny), Python (pandas, scikit-learn), SAS, C, C++, MATLAB

Interests

Statistical Genetics, Biostatistics, Biomedical Data Science, Health Informatics

RELEVANT EXPERIENCE

UNIVERSITY OF WASHINGTON

Independent Study Project

Advisor: Dr. Brian Shirts, Dr. Danny Miller

Seattle, WA

January, 2023 - present

- Analyzing the performance of various structural variant and SNV calling algorithms on long read sequencing data.
- Determining the accuracy of phasing algorithms using data from different trios.

ICAHN SCHOOL OF MOUNT SINAI

Department of Genetics and Genomic Sciences

Advisor: Dr. Paul O'Reilly

New York, NY

February – August, 2022

- Developed a novel shrinkage methods for improving prediction of a polygenic risk score model on UK Biobank data
- Summarized key algorithmic considerations and biological insights in the form of a manuscript to be submitted

COLUMBIA UNIVERSITY, GENENTECH INC.

Data Science Institute Scholar 2020-2021, Herbert Irving Center For Cancer Research

Advisor: Dr. Alison Taylor

New York, NY

October, 2020 – May, 2022

- Created a UNIX and R-based analysis pipeline on AWS consisting of copy number variant prediction algorithms on whole exome sequencing datasets to analyze differences between healthy and tumor samples.
- Quantified mutations (point and indels) stratified by chromosomal regions and genes across samples
- Working on a collaborative study with Genentech Inc. to understand correlations between immunotherapy status and copy number variations in multiple tumor types

INFLAMMATIX INC.

Computational biology summer intern

Advisor: Dr. Yehudit Hasin, Dr. Yudong He

Burlingame, CA

May – September, 2021

- Implemented a multi-cohort analysis framework on 20 publicly available datasets consisting of a heterogenous set of >2000 samples.
- Derived a gene expression signature to distinguish between healthy controls and Systemic Lupus Erythematosus (SLE), validated in an independent set with high accuracy.

INDIAN INSTITUTE OF TECHNOLOGY MADRAS

Dual Degree Thesis, Robert Bosch Center for Data Science and AI

Thesis advisor: Dr. Himanshu Sinha

Chennai, TN, India
July 2019 – May 2020

- Assessed multiple gestational age (GA) estimation models on preterm birth prediction for a cohort of North-Indian women.
- Designed a regression-based model for GA estimation in first trimester with highest preterm labelling accuracy [1].
- Implemented machine learning approaches using ultrasound metrics to predict GA accurately in second and third trimesters. [2]
- Worked in close collaboration with a team of clinicians at the Translational Health Sciences Institute, Delhi as a part of Bill and Melinda Gates foundation grant, Grand Challenges of India grant and BIRAC grant.
- Presented research at multiple Data Science and AI academic conferences.

STANFORD UNIVERSITY

Summer Research Intern, Department of Biochemistry

Advisor: Dr. Julia Salzman

Stanford, CA
May – August 2019

- Developed a generalized linear model using R to predict alternative splicing events consisting of 50 features using primate transcriptomic data combined with genomic information of key regulatory elements.
- Identified multiple SINE and LINE repeat aggregations conserved across six primate families.

STANFORD UNIVERSITY

Khorana summer fellowship intern, Department of Biochemistry

Advisor: Dr. Julia Salzman

Stanford, CA
May – August 2018

- Designed a bioinformatic analysis pipeline for analyzing circRNAs in Epstein Barr virus infected human B-cell transcriptomic datasets downloaded from SRA to identify differences in lytic and latent infection stages of the virus.
- Selected as one of the 40 Khorana fellowship grant recipients among 1000 applicants to conduct research at the Salzman Lab, Department of Biochemistry.

LEADERSHIP EXPERIENCE

UNIVERSITY OF WASHINGTON

Teaching Assistant, School of Public Health

Seattle, WA
September, 2022 - present

- PHRMRA 545: Statistical Topics for Biomedical Regulatory Affairs Professionals: Responsible for grading homework, holding office hours, attending lecture and providing additional help with statistical methods.
- BIOST 511: Medical Biometry: Responsible for grading homework, holding discussion sections to enable students to practice course concepts and weekly office hours. I am also responsible develop content for weekly discussion sections.
- PHG 200: Implications of Public Health Genomics for Society: Responsible for holding discussion sections comprising of debates based on articles related to course content. I also hold weekly office hours.

COLUMBIA UNIVERSITY

Teaching Assistant, Department of Biostatistics

New York, NY
August – December, 2021

- P8105: Data Science: Responsible for grading homework assignments and mentoring student groups for the final project. I also hold weekly office hour sessions that involve debugging and explaining tidyverse/dplyr functions to students.
- P8104: Probability: Responsible for conceptualizing questions for weekly assignments, holding office hours to clarify theoretical concepts, grading homework and examinations.

INDIAN INSTITUTE OF TECHNOLOGY MADRAS

Teaching Assistant

Chennai, TN, India
July, 2019 – May, 2020

- Quantitative and Population Genetics: formulated questions for quizzes and end semester examinations along with grading of 2 online research seminars for over 20 students
- Synthetic Biology: responsible for assessing 3 sets of research seminars for over 25 students

Founder and Chair, Biotech Research Club

September, 2017 – June, 2019

- Started a student organization dedicated at developing an interest in Synthetic biology among the students of IIT Madras where I was responsible for planning and conducting relevant talks by graduate students, experimental workshops to compliment theoretical understanding and journal clubs on dedicated topics.
- Curated over 10 articles for a science magazine called `Synkranti` as the chief editor

CONFERENCES

- Participant and scholarship recipient, 26th Summer Institute in Statistical Genetics, University of Washington (2021)
- Presenter, Initiative for Biological Systems Engineering 8th Workshop (2021)
- Presenter, 7th RBCDSAI Workshop on Recent Progress in Data Science and AI (2020)
- Presenter and volunteer, India EMBO symposium on Synthetic Biology (2020)
- Runners up and best software tool out of 15 teams, AI Genomics Hackathon, SVAI, San Francisco (2019)
- Initiative for Biological Systems Engineering 5th workshop (2019) – best poster and presentation
- Tata consultancy services research presentation competition – award for best poster and presentation (2019)
- International Genetically Engineered Machine Competition, Boston (2016-2018)
 - Gold medal, Foundational Advance track (2018)
 - Silver medal, Software track (2017) for published tool ChassiDex [3]
 - Silver medal, Measurement track (2016)

PUBLICATIONS

- [1] Vijayram R*, **Damaraju N***, Xavier A*, et al. Comparison of first trimester dating methods for gestational age estimation and their implication on preterm birth classification in a North Indian cohort. *BMC Pregnancy and Childbirth*. 2021; 21(1):343. doi:10.1186/s12884-021-03807-4 (*co- first authors)
- [2] **Damaraju N***, Xavier A*, et al. *Development of Second and Third-Trimester Population-Specific Machine Learning Pregnancy Dating Model (Garbhini-GA2) Derived from the GARBH-Ini Cohort in North India.*; 2021: 2021.10.02.21264450. doi:10.1101/2021.10.02.21264450 (*co- first authors)
- [3] Kailash BP*, Karthik D*, Shinde M*, **Damaraju N***, et al. *ChassiDex: A Microbial Database Useful for Synthetic Biology Applications.*; 2019:703033. doi:10.1101/703033 (*co- first authors)