Shreya Prakash

shreyap1@uw.edu

Research Interests: Causal Inference, Fairness, Sensitivity Analysis

Website: https://shreyap18.github.io/

EDUCATION

University of Washington 2020 - 2025 (expected)

PhD in Statistics, Advanced Data Science Track. *(Advisor: Elena Erosheva)*

Carnegie Mellon University 2016 - 2019

B.S in Statistics and Machine Learning, University Honors

RESEARCH EXPERIENCE

University of Washington

Dissertation Research Winter 2022 - Present

- Developing statistical methods for causal structure learning and causal decomposition

Research Assistant in WA Notify project Spring 2021-Fall 2021

- Advised by Debra Revere and Janet Baseman from the Department of Epidemiology
- Conducted research on the impact of the privacy-protected exposure notification app (WA Notify) on COVID-19 transmission and identified factors influencing willingness to quarantine and get tested using statistical methods

Carnegie Mellon University

Undergraduate Research Assistant advised by Alexandra Chouldechova Fall 2019

- Assessed the presence of age, race, or gender-based disparity in the utilization of fully or semi-automated decision-making processes for determining when a case worker should investigate specific abuse cases

Undergraduate Research Assistant advised by Peter Freeman Spring 2019

- Developed a data pipeline to aid astronomers in understanding the evolution of galaxies based on their current appearances, employing techniques to address imbalanced data

Undergraduate Research Intern in Black & Veatch Corporate Capstone Project Fall 2018 - Spring 2019

- Created an R Shiny app for analyzing historical company data, predicting injury and property damage cases, and generating prevention strategies through partial dependence plots (pdp).

Undergraduate Research Intern for the KONAM Foundation Fall 2017

- Designed and implemented a machine learning algorithm that assesses the risk of planting certain crops for marginalized farmers in India

PUBLICATIONS, CONFERENCE PRESENTATIONS, & MEDIA

1. **S. Prakash**, C. Cinelli, E. Erosheva, C. Lee, "A Causal Decomposition Analysis of Black-White Disparity in Selection into Discussion during NIH Proposal Review Process", (2024), (in preparation)
2. **S. Prakash**, F. Xia, E. Erosheva, "Towards Causal Discovery with Statistical Guarantees", (2024), (in preparation), Presented at The Western North American Region of The International Biometric Society 2023
3. **S. Prakash**, I. Javed, S. Lu, A. Adler, B. LeRoy, R. Nugent, "Characterizing Incidents at Black & Veatch", *Carnegie Mellon University Meeting of the Minds Undergraduate Research Symposium,* (2019)
4. **S. Prakash**, P. Freeman, "Linking Galaxies Across Time via Conditional Density Estimation", *Carnegie Mellon University Meeting of the Minds Undergraduate Research Symposium,* (2019).
5. S. Konam, **S. Prakash**, S. Papp, S. Mishra, X. Liu, Z. Ma, K. Siripurapu, "New App for Indigenous Farmers", *The Hans India,* (2017), Link.

PROFESSIONAL EXPERIENCE

Data Scientist, *Marinus Analytics* Winter-Summer 2020, Summer 2021

- Applied machine learning and time series analysis for unstructured child welfare case records
- Launched spam filter and underage person detection algorithms for TraffickJam: an application that uses human trafficking advertisement data to aid law enforcement with finding trafficking victims and traffickers
- Productionalized Infoshield: a text clustering algorithm for large scale human trafficking advertisement datasets.

Data Science and Research Intern, *84.51°* Summer 2019

- Fixed issues and tested optimization algorithms for grocery promotion; recommended running promotion optimization for 52 weeks to increase category performance by 4%

Software Developer Intern, *Optum Technologies* Summer 2018

- Built authentication, UI and containerized existing application for cloud deployment, generating millions in savings and revenue

Research Intern, *Royal Caliber D3M Program DARPA* Summer 2017

- Worked on machine learning on graph datasets and implemented a significantly more efficient way to estimate the number of triangles in a graph, (from $O(V^3)$ to $O(V)$, where V is the number of graph vertices), using a wedge sampling algorithm

SKILLS

Programming: Proficient in R, Python, SQL. Familiar with C, Standard ML, Matlab, Java, Bash, Mathematica

Libraries/Software: numpy, pandas, scipy, sklearn, statsmodels, tensorflow, seaborn, matplotlib, rjags, tidyverse, parallel, Git, Docker, AWS S3 & EC2, Keras

AWARDS

- Center for Statistics and the Social Sciences Travel Award 2023
- 3rd Place winner in Meeting of Minds Undergraduate Research Symposium Poster Presentation Competition 2019

TEACHING EXPERIENCE

Teaching Assistant, *University of Washington*

- Spring 2023: Quantitative Introductory Statistics for Data Science (STAT 391)
- Winter 2023: Elements of Statistical Methods (STAT 311)
- Autumn 2022: Statistical Reasoning (STAT 220)
- Spring 2022: Causal Modeling (STAT 566)
- Winter 2021: Statistical Concepts and Methods for the Social Sciences (STAT 221)

Teaching Assistant, *Carnegie Mellon University*

- Fall 2019: Introduction to Probability Theory (36-225)
- Spring 2019: Introduction to Machine Learning (10-601)
- Fall 2017 & 2018: Methods for Statistics and Data Science (36-202)

RELEVANT COURSEWORK

- Statistical Learning, Advanced Theory for Statistical Inference, Advanced Regression Methods, Causal Modeling, Foundations of Fairness in Machine Learning, Statistical Graphics and Visualization, Parallel and Sequential Data Structures and Algorithms (C/SML), Linear Algebra