Anav Sood

Email: anavsood@stanford.edu Phone: +1 (216) 406-3130

Education

Ph.D. in Statistics 2020 - 2025 (anticipated)
Stanford University
Advisor: Prof. Trevor Hastie

B.A. in Mathematics 2016 - 2020
M.S. in Statistics
Stanford University

Professional Experience

Data Science/Machine Learning Consultant 2024 - current
Coframe

- Consulting on problems involving using A/B testing and/or multi-arm bandits to compare LLM agents

Data Science Intern 2022
Wayve

- Developed framework and methodology for measuring correlation between vehicle performance in real world and simulation
- Developed new A/B tests (currently deployed) for comparing different self driving model's real world performance.

Data Science/Machine Learning Consultant 2020
General Catalyst - Customer Value Fund

- Quantified uncertainty surrounding customer churn in certain consumer businesses

Data Science/Machine Learning Consultant 2020
Sequoia Capital

- Designed procedure for automating talent search

Quantitative Research Intern 2019
Citadel

- Evaluated modern feature selection methods' efficacy in settings with high dimensional time series data
- Ran event studies to determine if bond rating changes were significant indicators of stock price movement

Software Engineering Intern 2018
Cruise Automation

- Invented patented algorithm for the naively NP-hard problem of identifying the k avoidance areas which most negatively impact routability

Research

Research interests: Interpretable machine learning; Adaptive statistical inference; Precision medicine.

First authorship indicated by *

Articles

3. Sood, A. and Hastie, T. (2024+) A Statistical View of Column Subset Selection . *Re-submitted after revision to the Journal of the Royal Statistical Society: Series B.* arXiv:2307.12892
2. Sood, A. (2024+) Selective inference is easier with p-values. *Submitted to the Annals of Statistics.* arXiv:2411.13764
1. Mayer, A. T.*, Holman, D. R.*, Sood, A.*, Tandon, U., Bhate, S. S., Bodapati, S., ..., and Rogalla, S. (2023). A tissue atlas of ulcerative colitis revealing evidence of sex-dependent differences in disease-driving inflammatory cell types and resistance to TNF inhibitor therapy. *Science Advances.* sciadv.add1166

Books

1. Sun, D., Kim, G., and Sood, A. (2024). The Art of Chance: A Beginner's Guide to Probability. *Preprint available here.*

Patents

1. Sood, A., Swofford, M., Rech, L. O. M., and Bowe, A. Analysis of network effects of avoidance areas on routing. U.S. Patent 10,962,380. Filed December 20, 2018. Issued March 30, 2021.

Talks

Selective inference is easier with p-values
International Seminar on Selective Inference February 2025
Art Owen's Group Meeting January 2025
Stanford Industrial Affiliates Conference November 2024
Stanford-Berkeley Joint Colloquium October 2024

A Statistical View of Column Subset Selection
STATS 305C: Applied Statistics III Lecture May 2024
Stanford Industrial Affiliates Conference November 2023
Stanford-Berkeley Joint Colloquium October 2023
Joint Statistical Meetings August 2023

Transformers, LLMs, and what statistics can offer
Statistical Learning Group Meeting February 2024

- Slides available here

Perspectives on Frequentism and Bayesianism
Stanford Statistics Department Retreat May 2024

- Slides available here

The modern dimensionality reduction toolkit
STATS 305C: Applied Statistics III May 2024

- Slides available here

Posters

Selective inference is easier with p-values
Statistics Empowering Data Science January 2025

- One of three poster award winners

A note on binary words avoiding given subwords
Joint Mathematics Meetings January 2018

- Paper in preparation

Software

- **pycss**, developer, https://github.com/AnavSood/CSS
  Python package for column subset selection methods developed in my dimensionality reduction work
- **seldom**, developer, https://github.com/AnavSood/seldom
  Python code for running experiments in my selective inference work

Service and Teaching Experience

Instructor

STATS 216V: Introduction to Statistical Learning Summer 2023
STATS 208: Bootstrap, Cross-Validation, and Sample Re-use Winter 2023

- Designed course material, available here, from scratch to accommodate remote learning

STATS 110: Statistical Methods in Engineering and the Physical Sciences Fall 2021

- Designed course material, available here, from scratch to accommodate remote learning

Teaching Assistant

STATS 216V: Introduction to Statistical Learning Summer 2024
STATS 305C: Applied Statistics III Spring 2024
STATS 116: Theory of Probability Autumn 2023
STATS 315B: Modern Applied Statistics: Learning II Spring 2022
STATS 100: Mathematics of Sports Spring 2021
STATS 207: Introduction to Time Series Analysis Fall 2020

- Recipient of Departmental Teaching Assistant Award for 2023-24

XCS224N: NLP with Deep Learning 2019-2020

- Member of teaching staff for Stanford Center for Professional Development's course XCS224N
- Designed and built all course assignments relating to transformers and double descent

Service

Member of Stanford Statistics PhD Admissions Committee 2024 - 2025

- Among first PhD students to join the committee

Member of Stanford Department of Music Search Committee for Director of Jazz Studies 2024 - 2025

References

Trevor Hastie
John A. Overdeck Professor of Mathematical Sciences
Professor of Statistics
Professor of Biomedical Data Science
Stanford University
hastie@stanford.edu

Robert Tibshirani
Professor of Statistics
Professor of Biomedical Data Science
Stanford University
tibs@stanford.edu

Stephan Rogalla
Clinical Assistant Professor, Medicine - Gastroenterology & Hepatology
Member, Maternal & Child Health Research Institute (MCHRI)
Stanford University
srogalla@stanford.edu

Jonathan Taylor
Professor of Statistics
Stanford University
jonathan.taylor@stanford.edu

Dennis Sun
Associate Professor (Teaching) of Statistics
Stanford University
dlsun@stanford.edu