Shamindra Shrotriya

Department of Statistics and Data Science, Carnegie Mellon University

☐ +1 415 937 3701 | ✉ shamindra.shrotriya@gmail.com | 🏠 shamindras.com | 🔲 shamindras

Education

Carnegie Mellon University	*Pittsburgh, PA*
Ph.D. in Statistics & Data Science	*Dec. 2022*
M.S. in Statistics & Data Science	*2019*
University of California, Berkeley	*Berkeley, CA*
M.A. in Statistics	*2016*
University of New South Wales	*Sydney, NSW*
BCom (Actuarial Science & Finance)	*2007*

Technical Skills

Coursework: Convex Optimization, Adv. Statistical Inference, Adv. Probability Theory, Deep Learning, Statistical Machine Learning, Data Structures and Algorithms.

Proficient: R, Python, `pandas/scikit-learn/Pyspark/numpy`, `PyTorch/Tensorflow/JAX`, `Git/Github`, SQL (`Hive/Redshift/PostgreSQL`), `Make`, SAS, `Unix`, `Julia`, `Docker`.

Industry Experience

Walmart, Retail Intelligence — *Bentonville, USA*
Principal Data Scientist — *2023 - current*

- Building flexible and accurate demand forecasting tools using Bayesian dynamic linear models.
- Developing the experimental design framework in improving our Customer Experience (CX) score.

Carnegie Mellon University — *Pittsburgh, USA*
Faculty, Co-Director (Optum Summer Bridges) — *2023 - 2023*

- Co-director of 8 week undergraduate summer research program with 30-35 students selected nationally.
- Co-taught daily data science lectures on advanced statistical methods using R and `Python`.
- Designed and taught a new data engineering course focused on advanced Unix and SQL concepts.

freelancer.com — *Sydney, Australia*
Team Lead, Data Science Infrastructure — *2014 - 2015*

- Designed and implemented a prototype of the new A/B testing framework.
- Co-designed and administered the Extract-Transform-Load (ETL) process written in Go and AWS Redshift.

Quantium Consulting — *Sydney, Australia*
Data Scientist — *2012 - 2014*

- Led the end-to-end development of the behavioural 'lifestage' customer classifier for the entire 7 million Woolworths Supermarket customerbase.
- Led the data-driven electronic marketing strategy for Woolworths Life Insurance which included developing scoring models (GLMs) and conducting A/B tests to optimise response rates.

United Nations - International Labor Organization — *Pune, India*
Microinsurance Fellow — *2011 - 2012*

- Wrote a report on the best actuarial pricing practices to be undertaken by microinsurance organisations.

PwC — *Sydney, Australia*
Senior Actuarial Consultant — *2007 - 2011*

- Built visualization dashboards for monitoring key risk metrics for Insurance Australia Group, Australias' largest private general insurer.

Publications

PAPERS

1. Bong, H., Li, W., Shrotriya, S., & Rinaldo, A. (2020). Nonparametric Estimation in the Dynamic Bradley-Terry Model. In *AISTATS (Online)*.

2. Li, W., Shrotriya, S., & Rinaldo, A. (2022). sup-norm Bounds of the MLE in the BTL Model under General Comparison Graphs. *Uncertainty in Artificial Intelligence (UAI)*.

3. Dalmasso, N., Shrotriya, S., & Reinhart, A. (2019). Predictive Inference of a Wildfire Risk Pipeline in the United States. *NeurIPS 2019 Workshop on Tackling Climate Change with Machine Learning*.

UNDER REVIEW (SUBMITTED)

1. Shrotriya, S., & Neykov, M. (2022). *Revisiting Le Cam's Equation: Exact Minimax Rates over Convex Density classes*.

2. Shrotriya, S., & Neykov, M. (2022). *Uniform Location Estimation on Convex Bodies*.

3. Shrotriya, S., & Neykov, M. (2022). *Adversarial Sign-Corrupted Isotonic Regression*.

4. Fogliato, R., Shrotriya, S., & Kuchibhotla, A. K. (2021). *maars: Tidy Inference under the "Models as Approximations" Framework in R*.

COMPETITIONS

1. Bong, H., Li, W., & Shrotriya, S. (2019). Efficient Estimation of Distribution-Free Dynamics in the Bradley-Terry Model. *Carnegie Mellon Sports Analytics Conference (Reproducible Research Winner)*.

2. Barter, R., & Shrotriya, S. (2016). Integrated Data Analysis for Early Warning of Lung Failure. *ODBMS.org (Geisinger Competition Winner)*.

Presentations

useR! 2021: The R Conference (Spotlight Talk)   *Virtual*

`maars`: Tidy inference under misspecified statistical models in R   *2021*

NeurIPS 2019 Climate Change Workshop (Spotlight Talk)   *Vancouver, BC*

Predictive inference of a wildfire risk pipeline in the United States   *2019*

Honors and Awards

rstudio::global(2021) Diversity Scholar (RStudio)   *2021*

NGC Wildfire Research Scholar (American Australian Association)   *2020*

TA of the Year (Carnegie Mellon University)   *2020*

CMSAC Best Paper Award (Carnegie Mellon University)   *2019*

NeurIPS Climate Change Workshop Travel Award   *2019*

Outstanding Graduate Student Instructor (University of California, Berkeley)   *2017*

Elizabeth Scott Memorial Award (University of California, Berkeley)   *2016*

Best Paper and Competition Winner (Geisinger Health Collider Project)   *2016*

Software

maars R Package   *2021*

Co-creator of the `R` package to implement tidy inference under the 'Models as Approximations' framework. Joint work with Riccardo Fogliato and Arun Kumar Kuchibhotla.

Iterative Random Forests (iRF) Python Package   *2017*

Co-Developer of `Python` package to detect predictive and stable high-order interactions.