

Ravi B. Sojitra

Curriculum Vitae

August 2020

United States citizen.

Native English speaker.

 RaviSoji.com

 RaviSoji [at] stanford [dot] edu

 RaviSoji

 RaviSoji

Education

- 2020– PhD **Stanford University**
Management Science and Engineering
Research Specialty: Computational Social Science
- 2018–2020 MS **New York University**
Applied Statistics for Social Science Research
Concentration: Computational Methods
GPA: 4.000/4.000
- 2010–2014 BA **Rutgers University—Newark**
Biology and Philosophy
Honors Thesis: Age Differences in Positive and Negative Probabilistic Feedback Learning
GPA: 3.954/4.000 (summa cum laude)

Overview of Relevant Work History

- 2019–2020 **Microsoft Research NYC**, Research Intern
Computational Social Science Group
Mentors: Drs. Shawndra Hill, David Rothschild, and K. Sudhir (Sudhir Karunakaran)
Projects: *Impact of (1) derogatory autosuggestions and (2) TV ads on household web search.*
- 2016–2020 **Rutgers University—Newark**, Research Technician
Department of Mathematics and Computer Science
Mentors: Drs. Patrick Shafto and Scott Cheng-Hsin Yang
Projects: *Explainable Artificial Intelligence, Cyber-Human Systems, and Human-Algorithm Interaction.*
- 2019–2019 **New York University**, Course Assistant and Grader
Program: Applied Statistics for Social Science Research
Professor: Dr. Ravi Shroff
Course: *Messy Data and Machine Learning.*
- 2012–2016 **Rutgers University—Newark**, Research Assistant
Center for Molecular and Behavioral Neuroscience
PI: Dr. Mark A. Gluck
Projects: *Effects of (1) aging, (2) exercise, and (3) genetics on striatal and hippocampal learning.*

Conference Talk

- Hill*, S, RB Sojitra*, D Rothschild, and K Sudhir (2020). "Measuring the Long Term Impact of TV Ads Using Search". Talk given at the 16th Symposium on Statistical Challenges in Electronic Commerce Research.

Research Articles

- Yang*, SCH, WK Vong*, RB Sojitra*, and P Shafto (2019). "Understanding explainable artificial intelligence with Bayesian Teaching". Under Review.
- Sojitra*, RB, I Lerner*, JR Petok, and MA Gluck (2018). "Age affects reinforcement learning through dopamine-based learning imbalance and high decision noise – not through Parkinsonian mechanisms". *Neurobiology of aging* 68, pp. 102–113.
- Vong*, WK, RB Sojitra*, A Reyes, SCH Yang, and P Shafto (2018). "Bayesian teaching of image categories". In: *Proceedings of the 40th Annual Conference of the Cognitive Science Society.*

Editorial

- Lerner, I, RB Sojitra, and MA Gluck (2018). "How age affects reinforcement learning". *Aging (Albany NY)* 1012, p. 3630.

Conference Posters

1. Richard, B, RB **Sojitra**, BC Hansen, and P Shafto (2018). "Defining non-linear processes in cross-orientation suppression (XOS) with steady state visual evoked potentials (SSEVPs)". Poster presented at the 18th Annual Meeting of the Vision Sciences Society.
2. **Sojitra***, **RB**, WK Vong *, and P Shafto (2017). "The Dynamics of Human Visual Experiences". Poster presented at the 1st Annual Conference on Cognitive Computational Neuroscience.
3. **Sojitra**, **RB**, C Inyang, A Shaw, and MA Gluck (2016). "Medicating hypertension may improve learning from positive probabilistic feedback, in African Americans". Poster presented at the 23rd Annual Meeting of the Cognitive Neuroscience Society.
4. **Sojitra**, **RB**, JR Simon, MA Gluck, and I Lerner (2016). "Reinforcement learning model reveals age group differences in cognitive strategies for probabilistic categorization". Poster Presented as the 38th Annual Meeting of the Cognitive Science Society (Neural Computation and Psychology Workshop).
5. **Sojitra**, **RB**, JR Petok, and MA Gluck (2014). "Adult age differences in "online" learning from positive and negative probabilistic feedback". Poster Presented as the 44th Annual Meeting of the Society for Neuroscience.

Open Source Code

- **Probabilistic Linear Discriminant Analysis:** <https://github.com/RaviSoji/plda>.
My Python implementation of the model presented by Sergey Ioffe (2006). I wrote it so that you can both (1) extract the features you would ordinarily want from plain linear discriminant analysis and (2) classify new data using the underlying probabilistic model. I also wrote unit, integration, and inference tests for sanity checks.
- **Utility Functions for Google Colaboratory:** https://github.com/RaviSoji/colab_utils.
If you work at an academic institution in the United States, you probably have unlimited Google Drive storage. `pull_from_gdrive()` and `push_to_gdrive()` are intended to be conceptually analogous to the pull and push commands in git.

Statistics and Modeling Skills

Causal Inference	Propensity score, diff-in-diff, instrumental variable, and regression discontinuity approaches. Multilevel panel data.
Regression	Binary, Count, and continuous treatments.
Classification	Linear (mixed effect) models, generalized linear (mixed effect) models; regression trees. Probabilistic, neural network, (binary and multinomial) logistic regression, and random forest approaches.
Sampling	GAN image sampling, collapsed Gibbs sampling, and other MCMC approaches.
Dimension Reduction	Principal component analysis, linear discriminant analysis, factor analysis. Fourier transform, discrete cosine transform, and autoencoding approaches.
Non-parametric Clustering	Dirichlet process mixture models (e.g. infinite Gaussian mixture models).
Parametric Clustering	k-Means and finite Gaussian mixture models.

Programming and Software Skills

Operating Systems	Unix (MacOS), Linux (CentOS and Ubuntu), Windows 10.
Software	Vim, Conda, Jupyter, Rstudio, tmux, Git, Slurm Workload Manager, Microsoft Office.
Languages	Python, R, Bash, ScopeScript/SQL, Markdown (and RMarkdown), LaTeX; C#.
Python packages	matplotlib, numpy, pandas, pytest, scipy, scikit-learn, unittest; pytorch, tensorflow.
R packages	Base R, tidyverse (e.g. dplyr, ggplot2, tidyr, readr, lubridate), models (e.g. glm, glmer, lm, lme4, lmerTest, randomForest, brms).

Statistics and Machine Learning Course Work

Causal Inference	Bayesian Statistics
Multilevel Models	Frequentist Statistics
Network Analysis	Multivariate Analysis
Experimental Design	Generalized Linear Models
Independent Study: Causality and Fairness	Multivariate Analysis
Independent Study: Heterogenous Treatment Effects	Probability
Supervised and Unsupervised Machine Learning	Linear Algebra
Messy Data and Machine Learning	Statistical Computing

Service to the Community, Organization, and Profession

- 2020 — Present **Co-organizer** 3+ hours/month — New York University.
Margarita Boyarskaya and I co-organize monthly reading group meetings on Causality and Fairness. Median attendance is approximately 10 people (4+ faculty) from various departments. Each meeting is 90 minutes long: 15 minute reading summary and 75 minute discussion. I typically moderate these discussions (in addition to participating).
- 2019 — 2020 **Panelist** 6-8 hours total — Microsoft Research NYC and New York University.
I participated in various panels for undergraduate students and prospective graduate school applicants. Most of these were organized for students from underrepresented backgrounds. At NYU, I also participated in subsequent small group meetings with students.
- 2017 — 2018 **Mentor** 5+ hours/week — Rutgers University, Newark campus.
Helped Professor Patrick Shafto supervise 1 undergraduate thesis project on Generative Adversarial Networks for image analysis and 1 MS thesis project on Dirichlet Process Mixture Models for video analysis.
- 2012 — 2016 **Co-organizer and Speaker** 10+ hours/week — African American Brain Health Initiative.
Prepared and delivered talks on brain health, Alzheimer's disease, and research participation (including legal rights and ethics) at senior centers, publicly assisted housing centers, and churches. Helped organize and participated in 15 outreach events in Newark & Greater Newark for audiences of 20-100.
Helped Post-Doctoral Fellows mentor 3 high school and college students to increase scientific literacy.
- 2010 — 2014 **Volunteer Tutor** 2+ hours/week — Rutgers University, Newark campus.
Subjects: Chemistry, Calculus I, Physics I & II, Ethics, Logic (including Modal logic).
Tutored approx. 10 students 1-on-1 for at least one semester each.
- 2012 — 2013 **Student Government Senator** 4+ hours/week — Rutgers University, Newark campus.
Won both Senate races I ran in.
Participated in a restructuring committee to more fairly represent various schools within Rutgers.
Worked to standardize ad-hoc student organization funding.
- 2010 — 2014 **Newark Volunteer (Miscellaneous)** 10+ hours/month — Newark, NJ, USA.
Volunteered in the emergency room at University Hospital 4-8 hrs/month, updating families in waiting rooms, talking to anxious patients, and transporting patients to tests and scans.
Participated in various one-off activities. E.g. painted buildings for non-profit organizations such as Kids Corporation and participated in Sleep Outs to raise awareness for hunger and housing issues.

Awards, Honors, and Scholarships

- 2020 Steinhardt ASH Leadership Award, New York University
- 2018-2020 21st Century Scholarship, New York University
- 2014 Charles I. Biederman Award (*Department award for Philosophy*), Rutgers University—Newark
- 2013 Phi Beta Kappa (*early induction*), Rutgers University—Newark
- 2010-2014 Dean's Scholarship, Rutgers University—Newark
- 2010-2014 Rutgers Scarlet Scholarship, Rutgers University—Newark
- 2010-2014 Newark Chancellor's Scholarship, Rutgers University—Newark