

# DEBJANI SAHA

Data Scientist, Research Statistician  
dsaha@cs.umd.edu | saharaja.github.io

---

## EDUCATION

- University of Maryland**, College Park, MD Aug. 2017 – May 2020  
M.S., Computer Science | Advisor: Dr. Michelle L. Mazurek
- New York University Tandon School of Engineering**, New York, NY Sept. 2016 – May 2017  
Additional computer science coursework (discrete math, algorithms)
- New York University**, New York, NY Sept. 2009 – May 2013  
B.S., Neural Science, Magna Cum Laude | Joint Minor, Computer Science & Mathematics | Pre-Medicine Track

---

## PROFESSIONAL EXPERIENCE

- Data Scientist** Jul. 2020 – present  
Ruxton Advisors, LLC; Sub-contractor for the Social Security Administration (SSA) Baltimore, MD
- Developing machine learning (ML) models for the SSA to predict likelihood of re-employment for Title II and XVI (disability) beneficiaries; using these models to inform future re-employment outreach efforts
  - Building a software package utilizing ensemble ML methods to impute missing data in the Occupational Requirements Survey (ORS) administered by the Bureau of Labor Statistics (BLS)
- Graduate Research Assistant** Sept. 2018 – Aug. 2020  
Maryland Cybersecurity Center, UMD College Park, MD
- User study (interviews) interrogating attitudes towards privacy and security concerns in commercial genetic testing
  - User study (surveys) evaluating lay comprehension of mathematical definitions of fairness used in machine learning
- Bioinformatics Specialist / Consultant** Nov. 2015 – Dec. 2019  
The Rockefeller University / Icahn School of Medicine at Mount Sinai New York, NY
- Built custom pipelines for analyzing gene expression data (microarrays, RNA sequencing) in immunological contexts
  - Worked to determine the effects of perturbations (e.g., infection, vaccination, and therapy) on human gene expression
- Research Associate** Oct. 2013 – Oct. 2015  
Lieber Institute for Brain Development, JHMI Baltimore, MD
- Built pipeline to automate extraction, preprocessing and analysis of data from functional neuroimaging (fMRI) studies
  - Worked to identify brain regions linked to both cognitive deficits in schizophrenia (SZ) and genetic inheritance

---

## SELECTED CONFERENCE PROCEEDINGS

- Saha, D.**, [and 4 others], & Mazurek, M.L. User attitudes on commercial genetic testing. Proceedings of 5<sup>th</sup> IEEE European Symposium on Security and Privacy (EuroS&P) 2020. Virtual Conference. [[Paper](#)]
- Saha, D.**, [and 4 others], & Tschantz, M.C. Measuring non-expert comprehension of machine learning fairness metrics. 37<sup>th</sup> International Conference on Machine Learning (ICML) 2020. Virtual Conference. [[Paper](#)]
- Saha, D.**, [and 4 others], & Tschantz, M.C. Human comprehension of fairness in machine learning. Proceedings of the 3<sup>rd</sup> AAAI/ACM Conference on AI, Ethics, and Society (AIES) 2020. New York, NY. [[Abstract](#)] [[ArXiv paper](#)]

## JOURNAL ARTICLES

- Seifert, L.L., Si, C., **Saha, D.**, [and 12 others], & Dittmann, M. The ETS transcription factor ELF1 regulates a broadly antiviral program distinct from the type I interferon response. PLOS Pathogens 2019, 15(11): e1007634. [[Article](#)]
- Wu, X., Thi, V.L.D., Huang, Y., Billerbeck, E., **Saha, D.**, [and 13 others], & Rice, C.M. Intrinsic immunity shapes viral resistance of stem cells. Cell 2018, 172(3): 423-438. [[Article](#)]
- Kaplan, C.M., **Saha, D.**, [and 5 others], & Tan, H. Estimating changing contexts in schizophrenia. Brain 2016, 139(7): 2082-2095. [[Article](#)]

---

## SELECTED SKILLS

- Programming** R, Matlab, Python, Perl, Java, Linux, shell (Bash) scripting, Git, LaTeX, [R] Markdown
- Specialized software** Tableau, Qualtrics, MaxQDA
- Quantitative methods** Experimental design, data analysis, statistical computing, predictive modeling/analytics
- Qualitative methods** Survey and interview design, user studies, grounded theory
- Human subjects research** Behavioral and biomedical settings, Institutional Review Board (IRB) procedures
- Scientific communication** Technical writing, literature research and review, data visualization
- Languages** Fluent in English, Bengali, basic German and Hindi; literate in English and German

---

## HONORS & OTHER EXPERIENCE

- Member, Phi Beta Kappa Society
- Martial arts (Taekwondo, Shotokan karate, BJJ) – two black belts with sixteen years of experience