

Steven Xu

6830 Crescent Moon Ct., Raleigh NC 27606 | (530) 601-3587 | sgxu@ncsu.edu | <https://stevengxu.github.io>

PROFESSIONAL SUMMARY

Final-year Ph.D. student studying Statistics. Self-motivated researcher on **Bayesian Non-parametrics, Causal Inference, and Anomaly Detection**. Statistician with 2+ years industry experience in Machine Learning modeling.

EDUCATION

North Carolina State University, Raleigh, NC
Doctor of Philosophy: Statistics
Member of American Statistical Association

Expected Spring 2022

University of Washington, Seattle, WA
Bachelor of Science: Statistics
Minors: Applied Mathematics, Mathematics
Dean's list: 2015 - 2017

June 2017

SKILL

R, Python (NumPy, pandas, Matplotlib, scikit-learn, PyTorch, TensorFlow), SQL, Spark, SAS, Julia, Git, Hadoop

WORK EXPERIENCE

Wells Fargo, Remote

Summer 2021

Data Scientist Intern, AI Model Development

- Developed a ML classifier to detect human login attempts from blacklisted IPs in near real-time.
- Analyzed human/bot login patterns from historical records using linear/nonlinear embedding techniques.
- Researched on semi-supervised learning in presence of class imbalance (2% event rate) and noisy labels.
- Proposed a flexible and interpretable modeling solution that improves primary business metric by 20%.

SAS Institute, Cary, NC

Summer 2019 – Present

Machine Learning Intern, Internet of Things

- Develop online anomaly detection algorithms to monitor high-frequency sensor data in IoT applications.
- Contribute to DevOps of features and plug-ins to support the ESP (Event Stream Processing) platform.
- Routinely present research findings to technical and non-technical audiences.

UNC Center for AIDS Research, Chapel Hill, NC

Summer 2019

Quantitative Intern, Biometrics

- Studied the effect of an HIV antiviral on suicidal tendency using rare-event modeling.
- Researched on uncertainty analysis of causal estimators under missing covariates.
- Quantified the attenuation of causal effects when trial and target population significantly differ.

PROJECT EXPERIENCE

Real-time Signal Decomposition using Independent Component Analysis

Fall 2020 – Present

- Researched on independent component analysis (ICA) for blind source separation of streaming data.
- Applied ICA to extract and identify distinct components from image and speech data.
- Prototyped an online ICA algorithm to be integrated in SAS stream analytics platform.

Anomaly Detection using Sparse and Low-rank Matrix Decomposition

Fall 2019 – Fall 2020

- Researched on robust PCA for detecting image outliers in video surveillance and remote sensing applications.
- Applied robust PCA and connected-component analysis to detect moving objects in stable background.
- Developed unsupervised wildfire detection and cloud masking algorithms for multispectral satellite imagery.

Bayesian Non-parametric Quantile Process Regression

Spring 2019 – Present

- Proposed a novel method for simultaneous estimation of non-crossing, non-linear quantile curves.
- Modeled high-dimensional conditional outcome distribution using Bayesian Neural Network.
- Conducted Bayesian causal inference on outcome quantiles in the presence of many confounders.
- Implemented model agnostics tool to extract interpretability from black-box quantile regression model.

PUBLICATION

Xu, S. G., & Reich, B. J. (2021). Bayesian Non-parametric Quantile Process Regression and Estimation of Marginal Quantile Effects, In press. *Biometrics*.

Xu, S. G., Kong, S., & Asgharzadeh, Z. (2021). Wildfire Detection Using Streaming Satellite Imagery. *Proceedings of the 2021 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*.

Mollan, K. R., Pence, B. W., **Xu, S.**, Edwards, J. K., Mathews, W. C., O'Cleirigh, C., ... & Bengtson, A. M. (2021). Transportability from Randomized Trials to Clinical Care: On Initial HIV Treatment with Efavirenz and Suicidal Thoughts or Behaviors. *American Journal of Epidemiology*.