Zihao (David) Xu

zxu@g.harvard.edu
https://github.com/zihaoxu/My_Portfolio (310) 962-6992

EDUCATION

Harvard University **Cambridge, MA**
Master of Science in Data Science *December 2020*

- Related coursework: Data Science I, Systems Development for Computational Science, Machine Learning

Pomona College **Claremont, CA**
Bachelor of Arts, Computer Science & Mathematics *May 2019*

- GPA: 3.96/4.0, Magna Cum Laude, Phi Beta Kappa Member, Sigma Xi Member, Pomona Scholar
- Related coursework: Artificial Intelligence, Algorithms, Computational Statistics, Big Data, Stochastic Operations Research, Probability, Statistical Theory, Bayesian Statistics, Econometrics

WORK EXPERIENCE

Cornerstone Research **Los Angeles, CA**
Summer Analyst *Summer 2018*

- Cleaned, managed, and visualized datasets with >10 million observations using R, SQL and SAS to advise two major auto manufacturers on recall strategies
- Interpreted and critically analyzed the validity of statistical sampling methods employed by opposing expert
- Led R workshops for full-time analysts and enforced data organization and workflow within case teams
- Received full-time return offer in recognition of work quality and work ethics

Scratch Financial, LLC **Pasadena, CA**
Business and Credit Risk Analyst Intern *Fall 2017 – Fall 2018*

- Performed A/B testing on two types of signage sent to 400 vet clinics and identified the signage that yielded statistically significant improvement of 11% in the average number of daily applications
- Devised and selected 6 out of 35 new variables that filter out high-risk loan applicants and reduced company overall default rate by 2 percentage points
- Built an actively used dashboard that visualizes temporal and geographical trends in important metrics
- Engineered scalable pipelines to clean and wrangle previously unhandled consumer credit data

Pomona College Mathematics Department **Claremont, CA**
Summer Researcher *Summer - Fall 2017*

- Devised and implemented a new machine learning algorithm - Bag of Little Random Forests (BLRF) as an R package, utilizing a paralleled structure for faster computation
- Evaluated and visualized the statistical and computational performance of BLRF using simulated data sets
- Fine-tuned hyper-parameters of BLRF and analyzed their impact on accuracy and time consumption
- Authored paper, *Bag of Little Random Forests,* and won the first prize at 2017 Fall USRESP Competition

Claremont Chinese Student Association (CSA) **Claremont, CA**
President *Fall 2016 – Fall 2017*

- Planned and coordinated 5 major on-campus events to promote awareness of Chinese culture
- Acquired $4500 as external funding through successful negotiations with outside partners
- Provided 450 members with career assistance through alumni connections and alumni on-campus lectures

SKILLS & MEMBERSHIP

Tech Skills: Python (pandas, numpy, matplotlib, seaborn, sklearn, keras), R, SQL, C, A/B Testing, Web Scraping, linux, version control (git), Machine Learning, Computer Vision

Memberships: Quantitative Researcher @ Bahnaha Capital, Technical Lead in Data Science @ Tech for Good

Publications: **Z. Xu**, M. Salloum (2018), Deep Neural Networks for Object Enumeration, poster paper & poster presentation, 2018 IEEE International Conference on Big Data

Z. Xu, J. Hardin (2017), Bag of Little Random Forests (BLRF), First Prize, 2017 Fall USRESP Competition, available at https://bit.ly/2kecdFE