Xiang Li

PhD Candidate (3rd Year)
Statistics Major
The George Washington University

"Pray for what you want, work for what you need"

Mob.: +1-6263623428
Email.:xiangli94@hotmail.com

Education

2018-2022
PhD, Statistics, GPA 3.95/4.00
The George Washington University

2016-2018
Master, Statistics
The George Washington University

2015-2016
Exchange Program
Mathematics
Polytecnic University of Catalonia, Spain

2012-2016
BS, Applied Mathematics
Tongji University, China

Skills

Programming
R, Rcpp, LaTeX(Advanced)
Python (Intermediate)
SAS, Shell, C++, SQL (Beginner)

Language
Mandarin(Native)
English(Proficient)
Spanish(Elementary)

Coursework

Advanced Statistical Theory
Applied Linear Model
Applied Multivariate Analysis
Bayesian Statistics
Categorical Data Analysis
Database Technology Applications
Econometrics
High Dimensional Data Analysis
Mathematical Statistics
Non-parametric Inference
Probability
Statistical Data Mining
Statistical Computing
Web Technology and Applications

Interests

Watching Documentaries
Collecting Earrings (300+)
Building Lego

Work/Research Experience

2020.05 - 2020.08 Food and Drug Administration   ORISE Fellow
● Applied machine learning method support vector machine(SVM) and its variants (penalized SVMs, RFE-SVM) to conduct feature selection, to detect important factors that drive opioid prescription in US.
● Proposed and implemented a new standardization method that helps penalized regression models to select variables fairly.
● Coded simulations and wrote manuscript.

R, Python, Pandas, Numpy, Scikit-learn

2019.01 - 2019.08 Milken Institute School of Public Health

Research Assistant
● Processed and analyzed multiple data sets of dementia study, fitted GLM, GLMM and other statistical models to assess predictors effects on cognitive ability. Implemented the idea of pooling data sets with privacy restrictions by using simulation and causal inference.
● Studied the mid-to late-life blood pressure trajectory of subgroups in a cohort, fitted statistical models, and visualized the plots.
● One paper "Proof of concept example for use of simulation to allow data pooling despite privacy restrictions" is under review in journal Epidemiology.

R, plotly, ggplot2

2019.08 - PRESENT Dissertation Research
● Conduct community detection/cluster analysis in network data.
● Propose a new model framework with latent variables structure, which can apply to raw data instead of the binary adjacency matrix commonly employed by traditional cluster analysis model. Current research work involves coding stochastic gradient descent algorithms, MCMC, and variational EM algorithm to find optimal values in derived optimization problem.

R, Rcpp(C++), LaTeX

2018.01 - PRESENT The George Washington University Teaching Assistant
● Assisted in courses: GLM in Biostatistics(PhD), Mathematical Statistics(Grad), Contingency Table Analysis(Grad), Bayesian Analysis(Undergrad), Intro to Statistics in Social Science(Undergrad)
● Teach in recitations for undergraduates and hold office hour every week.

Leadership

2017.09 - PRESENT Leader, R Programming Club in Statistics Department
● Organize and give lectures in club meetings every three weeks. Topics cover introduction to some basic and popular R packages.

2018.01 - PRESENT Organize Workshops
● Co-organized " Data Challenge D.C." Hackathon with George Mason University statistics department, attended by 30 students.
● Organized and hosted workshops "PhD Application Panel", and "US Workplace Culture & Etiquette" etc, attended by 100 students in total.