

LINDA ZHOU

lzhou54@jhu.edu | (607)-3798051 | [LinkedIn](#) | [GitHub](#)

WORK EXPERIENCE

Data Scientist - Biostatistician

Jun 2020 - Present

Johns Hopkins Bloomberg School of Public Health

Baltimore, MD

- Lead data scientist for 10+ data science projects. Leveraged 10000+ patients' clinical data and 5000+ multi-omics (genomics, proteomics, metabolomics) data, used sophisticated statistical techniques (linear mixed model, random forest, clustering analysis, etc.) to investigate predictors of adverse kidney outcome.
- Managed Johns Hopkins Hospital clinical database for the Division of Nephrology using SQL, queried datasets for clinicians and researchers and provided guidance on appropriate usage of data.

Data and Policy Analyst Internship

Jun 2019 - Aug 2019

Acumen, LLC

Burlingame, CA

- Applied sophisticated statistical methodologies to compare demographic characteristics, claim information congruity and vaccination rate of MA beneficiaries with other Medicare beneficiaries.
- Independently developed analytical web applications using R shiny for interactive visualization of spatial data.

SELECTED PROJECTS

Sudden Cardiac Death Risk Prediction

- Implemented novel time-varying random forest models (RF-SLAM) on time-dependent variables and complex interactions, increased sudden cardiac death risk prediction AUC from 0.80 to 0.89 compared to Poisson regression model.

NBA Playoff Team Prediction

- Built machine learning models (Bradley-Terry Logistic Regression Model, Random Forest) that predicted the winning percentages of the NBA teams in the Western Conference (the playoff teams in 2020), achieved an overall accuracy of 87.5%.

Protein Function Prediction with Clustering Analysis

- Leveraged 5000+ proteomics data and 1000+ metabolomics data, selected potential proteins/metabolites that are associated with adverse kidney outcome with novel clustering analysis algorithm (Netboost).

PROFESSIONAL SKILLS

- Programing languages: R, Python, Bash, SAS, SQL, LaTeX
- Professional skills: machine learning, data management, data science, data visualization, AB testing, statistical consulting, genetics

EDUCATION

Johns Hopkins Bloomberg School of Public Health, Baltimore, MD

ScM Biostatistics, Class of 2020

Relevant coursework: Advanced data science I-II, Causal Inference, Introduction to Data Management, Method in Biostatistics I-IV, Probability and Statistical Inference I-IV, Statistical Computing

Cornell University, Ithaca, NY

B.S. Biometry and Statistics & Agricultural Sciences, Class of 2018

Relevant coursework: Categorical Data, Introduction to Computing with Python, Linear Models with Matrices, Linear Algebra, Multivariable Calculus, Probability Models & Inference, Theory of Statistics