Yi Han

75 Peterborough St, Boston, 02215 | 607-327-7183
yhan@hsph.harvard.edu | yiiihan.github.io

EDUCATION

Shanghai Jiaotong University (SJTU), Shanghai, China Sept. 2019 - Jun. 2023 (Expected)
Bachelor of Science in Statistics GPA: 3.93/4.3(90.3/100) Rank: Top 2

- **Relevant Coursework:** Probability Theory (95/100), Multivariate Statistics (96/100), Stochastic Processes (99/100), Optimization (90/100), Stochastic Simulation (92/100), Mathematical Analysis (95/100), Numerical Analysis (92/100), Math Finance (93/100), PDEs (93/100), ODEs (91/100), Complex Analysis (96/100)
- **Major Honors:** Fan Hsu-chi Scholars (awarded to 10 undergraduates each year), Outstanding student (top 5%)

Cornell University, Ithaca, NY Jan. 2022 - May. 2022
Exchange Student, Department of Statistics and Data Science GPA: 3.952/4.3

- **Relevant Coursework:** Machine Learning (A), Time Series Analysis (A), Undergraduate Study in Statistics (A+)

PUBLICATIONS

[1] Gu T, **Han Y**, Duan R.(2022) A transfer learning approach based on random forest with application to breast cancer prediction in underrepresented populations *Proceedings of Pacific Symposium on Biocomputing* [LINK]

[2] Gu T, **Han Y**, Duan R.(2022) Robust angle-based transfer learning built on ridge regression *ArXiv* [LINK]

EXPERIENCE

Research Assistant **Department of Biostatistics, Harvard T.H.Chan School of Public Health**
Supervised by Professor Rui Duan Jun. 2021 - Present

- Exploit transfer learning methods based on machine learning algorithms and application on biomedical datasets
  - Proposed a random forest-based TL framework targeting risk prediction and improved AUC substantially
  - Designed an angle-based transfer learning framework and conducted cross-validation to select tuning parameters
  - Reproduced existing DNN and CNN transfer learning algorithms with Pytorch to biomedical data.
  - Developed Rpackage **multiTL** for multiple transfer learning methods [LINK]
- Proposed a ranking-based polygenic scores ensemble method, and generated PRSs using multiple summary statistics and PRS estimation method (LDpred, PRSCS, SDPR, Sbayes, Lassosum)
- Conducted statistical analysis using Wilcoxon signed rank test and GLMM on survey data in Guinea Epilepsy Project.
- Demonstrated the genetic diversity of Hispanic ethnicity based on PCA of genome-wide SNP data.

How does mask mandate effect online learning? From a regression discontinuity perspective **Cornell University**
Independent Study Advised by Professor Yang Ning Feb. 2022 - May 2022

- Constructed a regression discontinuity framework including sharp regression with discontinuity design and constant treatment effect model and discovered negative treatment effects on the face mask mandate had on online learning.
- Evaluated the influence of the COVID-19 pandemic on education inequity by comparing the effects of face masks across school districts varying in percentage of minorities and per pupil total expenditure.

Selecting Hyper-parameters for Options Pricing Model **Financial Engineering Research Center, SJTU**
Independent Study Advised by Professor Yingda Song Jun. 2021 - Mar. 2022

- Conducted simulations on five continuous Markov chain grids in the pricing of European double barrier options.
- Analyzed the applicability and features of these grid design methods under the choice of underlying asset model.
- Designed adaptive grids by iterating continuous Markov chain to simulate strike prices at expiration using Monte Carlo.

Multi-factor Stock Selection Model Based on Regression Model **MS&E, Stanford University**
Independent Study Advised by Dr. Chenru Liu Jan. 2021 - Apr. 2021

- Preprocessed stock data to construct market and financial indicators and validate factors using single-factor analysis.
- Performed PCA tests to select factors and reduce multicollinearity.

Data Science Intern **Shanghai Fields Technology**
Intern at the Data & Algorithm Team Jul. 2021 - Aug. 2021

- Performed front-end configuration to regional carbon rating related data using SQL in investment advisory system.
- Utilized Python module to crawl statistical contents on web pages.

Sales & Trading Intern **Huaying Securities**
Intern at the Sales & Trading Department Jan. 2021 - Feb. 2021

- Evaluated changes in interest rates, coupon rates, maturity and bond credit ratings for a bond issuance report.
- Contributed to qualifications investigation reports for issuing corporate bonds.

ADDITIONAL INFORMATION

- **Leadership:** President of Student Union, School of Mathematics, SJTU
- **Programming languages:** R, Python (Pytorch, NumPy, Pandas, Matplotlib, Seaborn), Matlab