Hanyu (Mason) Liu

Ph.D. Candidate | Data Science / Applied Scientist | linkedin.com/in/hliu83

EDUCATION

Ph.D. in Geography and Environmental Health and Engineering 2023 - Present
Johns Hopkins University, Baltimore, MD. Greenhouse Gas Research Lab | Advisor: Dr. Scot Miller

Master of Health Science (MHS) in Environmental Health and Engineering 2022 - 2023
Johns Hopkins University, Baltimore, MD. Award: Outstanding Master's Essay | GPA: *4.0/4.0*

Bachelor of Science in Business Analytics, Business Analytics; Communications 2018 - 2022
George Washington University, Washington, D.C. Magna cum laude | GPA: *3.84/4.0*

EXPERIENCE

THE JOHNS HOPKINS UNIVERSITY Baltimore, Maryland
Greenhouse Gas Research Group – Applied Data Scientist / Researcher July 2022 – Present

- Built large-scale statistical modeling pipelines integrating 500,000+ observations from aircraft and in-situ tower sites across North America to evaluate state-of-the-art process-based methane flux models
- Applied inverse modeling, regression, and uncertainty quantification to identify systematic bias in predictive models, detecting ~2x overestimation in key regions such as Canadian boreal ecosystems
- Designed data workflows on Linux/HPC environments, processing high-dimensional matrices and time-series data using Python, R, and MATLAB

FI CONSULTING Washington, DC
Data Science & Business Analytics Consultant January 2022 –May 2022

- Developed and evaluated supervised machine learning models (Decision Trees and Support Vector Machines (SVM)) to predict financial distress and closure risk across 10,000+ institutional records
- Conducted model performance evaluation using metrics such as accuracy, precision–recall, and cross-validation, achieving a 15% improvement in predictive performance over baseline approaches
- Communicated technical findings to senior consultants and non-technical stakeholders via decision-oriented visualizations and summaries, informing risk-assessment and strategic planning decisions

PEKING UNION MEDICAL COLLEGE HOSPITAL Beijing, China
Medical Research Assistant/Data Analyst at Dr. Kaifeng Xu's Lab January 2021 – August 2021

- Analyzed high-dimensional clinical datasets using multivariate logistic regression, hypothesis testing, and retrospective cohort analysis to identify risk factors associated with rare pulmonary diseases
- Processed and validated patient-level data containing 30+ clinical, demographic, and biomarker variables, performing missing-data handling, outlier screening, and feature normalization
- Designed and executed case-control analyses with confounder adjustment, reporting odds ratios and 95% confidence intervals to quantify associations between biological indicators and clinical outcomes

TECHNICAL SKILLS

Programming & Data: Python, R, MATLAB, SQL, JMP Pro
Statistical & ML Methods: Logistic & linear regression, supervised ML, time-series analysis
Model Evaluation: Cross-validation, precision/recall, confidence intervals, feature importance
Systems & Tools: Linux, HPC workflows, AWS (compute & storage), Tableau
Software & Applied: Swift (iOS app development), web applications, AI-assisted prototyping

Languages: Chinese (Native), English (Fluent), Spanish & French (Elementary Proficiency)
Work Authorization: U.S. Permanent Resident (no sponsorship required)