

EDUCATION**National University of Singapore (NUS)****Aug 2020 – Jun 2021**

Master of Science in Statistics (CAP: 4.85/5.00, top 10/200)

- Coursework: Advanced Statistics I & II (PhD courses), Advanced Topics in Applied Statistics (PhD course), Analysis of Complex Biomedical Data Analysis (PhD course), Statistical Models: Theory and Applications (PhD course), Survival Analysis, etc.

Nanjing University (NJU)**Sep 2016 - Jul 2020**

Bachelor of Science in Mathematics (GPA: 4.30/5.00)

- Coursework: Mathematical Analysis, Data Structures, Statistical Analysis with Genetic Data, etc.
- Exchange Program with University of Queensland in Australia (GPA: 6.25/7.00)

RESEARCH INTERESTS

Deep learning, XAI, electronic health records, statistical genomics.

RESEARCH AND WORK EXPERIENCE**Senior Research Assistant****Jul 2021 – Present**

Duke-NUS Medical School (supervised by Prof. Nan Liu)

- Apply deep learning methods to PAROS registry data with size of ~200,000 for OHCA (out-of-hospital analysis) study, dealing with issue of data heterogeneity. Write paper as co-first author.
- Research on interpretable machine learning and build SHAP (SHapley Additive exPlanations) models on MIMIC-III data, dealing with data imbalance issue and proposing a weighted version. Write paper as first author.
- Conduct systematic review on missing data imputation in clinical area and write paper as co-first author.
- Take lead in systematic review on XAI in healthcare.

Research Assistant Intern**Apr 2021 – Jul 2021**

A*star – Agency for Science, Technology and Research (supervised by Dr. Chin Su)

- Conducted machine learning models (e.g., Random Forest and XGboost) via R and python to predict protein production based on frequency of amino acids, improving the general performance by 12% and writing pipelines to make the procedure efficient.

Data Science Intern**Nov 2020 – Apr 2021**

Benshi.ai

- Analyzed in project about maternal health and dealt with big datasets via Python and Spark through Azure cloud platform, providing efficient data analysis pipeline for modelling.
- Applied survival models via Pysurvival module to predict users' active time for Safe Delivery App and explored users' profiles, generating data insights with team members.
- Managed corporation cross teams with development of Machine Learning models and promote friendly cooperation with Denmark company Maternity Foundation.

Research Volunteer**Nov 2019 – Dec 2019**

University of Queensland (UQ)

- Researched and executed literature review on LD-score regression and GWAS, acquiring knowledge and skills such as linkage disequilibrium and sampling methods such as jack-knife.

Research Assistant Intern**May 2019 - Jul 2019**

Nanjing University (NJU)

- Trained convolutional neural network models via Python and TensorFlow based on data provided by Jiangsu Provincial People's Hospital, helping doctors to arrange tight beds more reasonably and highly improve quality of medical care.

STUDENT PROJECTS AND CONTESTS

Genome-wide Association Study Analysis and Complex Traits Analysis

Jun 2019 - Dec 2019

University of Queensland (UQ)

- Conducted genome-wide association analysis via R and Plink on genetic and phenotypic samples of nearly 11 thousand individuals.
- Applied biostatistical methods of REML and BLUP via R and GCTA revealing genetic architectures of three particular phenotypes about sleep duration and finally obtaining a score of distinction for project.

Mathematical Contest on Modelling (MCM)

Jan 2019 - Jan 2019

Consortium for Mathematics and Its Applications, COMAP

- Executed data cleaning and visualization via R and Tableau on data from U.S. Census Bureau and constructed models based on adjacency matrix applying Floyd Algorithm, PCA and influential network models, described characteristics and spread of specific synthetic opioid and heroin incidents.

Contemporary Undergraduate Mathematical Contest in Modelling

Oct 2018 - Oct 2018

China Society for Industrial and Applied Mathematics, CSIAM

- Led group and delegated tasks to establish models based on graph theory using Monte-Carlo method and dual programming via Python, developed evaluation on attained results.
- Co-wrote article achieving second price in national finals (first 5%).

EXTRA-CURRICULAR ACTIVITIES

Excellent Volunteer, Nanjing University

Nov 2018 - Nov 2018

- Guided a patient with cerebral palsy to type with computer and kept in touch with patient's family, completing patient's dream of expressing thoughts in words.

Vice President, Campus Ambassador Team - Teach for China

Oct 2017 - Oct 2018

- Organized and participated in activities of "Tutors Accompanying Reading", studying together and sharing ideas with children in remote mountain villages of Yunnan Province via letters and phone calls for one semester and received good feedbacks from students.

PUBLICATIONS

- Liu N, Liu M, Chen X, et al. Development and validation of an interpretable prehospital return of spontaneous circulation (P-ROSC) score for patients with out-of-hospital cardiac arrest using machine learning: A retrospective study[J]. *eClinicalMedicine*, 2022, 48: 101422.
- Liu M, Yuan H, Liu N. *Explain deep learning with SHAP towards imbalance: an empirical study on effects of background data and explanation data*. In progress.
- Liu M, Li S, Yuan H, Liu N. *Deep Learning for missing value imputation healthcare data: A systematic review of methodology*. In progress.

CERTIFICATIONS

- *Deep Learning Specialization* (5 courses) from DeepLearning.ai and Coursera
- *Big Data Fundamentals with Pyspark* from DataCamp
- *Finding Hidden Messages in DNA* (with Honors) from the University of San Diego and Coursera
- *Design and Interpretation of Clinical Trials* from Johns Hopkins University and Coursera
- *Introduction to Healthcare* from Stanford University and Coursera

SKILLS AND LANGUAGES

- IT / Software:
Proficient in R and Python, familiar with SQL, Spark and Tableau, basic in TensorFlow and Torch
- Language Skills: Fluent in English and Mandarin (both spoken and written)