

Miao Di

✉ di7miao@gmail.com 🏠 Homepage 

EDUCATION

NATIONAL UNIVERSITY OF SINGAPORE

Aug 2021 - Jan 2023

Master of Science in Statistics

Graduate Certificate in *Data Mining for Industry (Competence)* obtained as part of the Master's program.
With relevant coursework in *CS5288 Knowledge Discovery and Data Mining*, *ST5227 Applied Data Mining*.

BEIJING NORMAL UNIVERSITY

Sep 2017 - Jul 2021

Bachelor of Science in Statistics

EXPERIENCE

DUKE-NUS MEDICAL SCHOOL - Singapore

Jan 2023 - Jan 2025

A research-focused institution in collaboration with Duke University and the National University of Singapore.

Senior Research Assistant

- Developed and applied machine learning and AI models using R and Python to analyze clinical data, including analyzing datasets, optimizing algorithms, assessing model performance, and interpreting results to provide actionable insights.
- Participated in various research projects involving data processing and analysis. Collaborated with multidisciplinary teams to design experiments, develop methodologies, and interpret data to advance research objectives.
- Contributed to several peer-reviewed articles on machine learning applications in clinical and biological data analysis, including conducting literature reviews, performing detailed data analysis, and assisting in manuscript drafting.

ENVISION DIGITAL - Singapore

Mar 2022 - Jul 2022

Data Analyst

- Analyzed real data about solar energy using Python and SQL to identify patterns and trends.
- Generated reports based on data analysis findings to inform project decision-making.
- Collaborated with team members to improve programming for more accurate data pattern identification.
- Contributed to enhanced performance in troubleshooting and forecasting through data analysis insights.

ACADEMIC PROJECTS

Federated Learning for Clinical Structured Data:

A Benchmark Comparison of Engineering and Statistical Approaches

Nov/2023

- Key contributor to the Federated Learning (FL) benchmark project, a distributed AI approach that enables collaborative model training across decentralized data sources while preserving data privacy. Focused on coding and executing FL algorithms in R and Python, optimizing their performance, and conducting detailed evaluations. Created visualizations to present results effectively and provided crucial input to enhance the research paper. Contributed significantly to formal analysis and methodology design, and played a key role in writing and editing tasks.

FairFML: Fair Federated Machine Learning with a Case Study on

Reducing Gender Disparities in Cardiac Arrest Outcome Prediction

Oct/2024

- Played a key role in the FairFML project by generating visualizations (e.g., ROC curves, fairness metrics) to effectively communicate experimental results, enhancing the clarity and impact of the research findings. Organized and summarized experimental data, contributing to the refinement of the methodology and discussion sections in the research paper. Collaborated with the team to validate the framework's performance in reducing gender disparities while maintaining predictive accuracy, ensuring robust results for real-world clinical applications.

ADDITIONAL

- Technical Skills: R, python, SQL, SAS, Power BI, LaTeX
- Additional Courses: Completed *Managing Big Data with MySQL* course on Coursera, with certificate.
- Language Skills: English (Business Professional), Mandarin (Native), Japanese (Conversational)
- Certifications: Japanese-Language Proficiency Test (JLPT) Level N2