Jingxuan WU

Phone: (+86) 199-0656-6789 | Email: jingxuanwu@link.cuhk.edu.cn

EDUCATION

The Chinese University of Hong Kong, Shenzhen (CUHKSZ) Shenzhen, China
B.S. in Data Science and Big Data Technology Sept 2021 - Jun 2025

- **GPA:** 3.62/4.00
- **Relevant Courses:** Python, Optimization, Data Structure, Probability, Statistics, Econometrics, Machine Learning, Stochastic Process, Numerical Methods, Stochastic Simulation, Operations Management, Advanced Machine Learning
- **Honors:** Dean's List (2021-2023), Undergraduate Research Assistantship (2023,2024)

RESEARCH EXPERIENCE

Resource Allocation for Machine Learning Tasks Chapel Hill, NC
Supervisor: Prof. Guanting Chen, Dept. of Stat. & Operations Research, UNC May 2024 - Present

- Aimed to use large language models (LLM) to predict the runtime and energy consumption of machine learning tasks on different GPUs and achieve resource allocation optimization
- Extracted features from the code and applied LLMs to convert the code into embedding vectors suitable for prediction
- Trained Linear Probe to predict task execution time and energy consumption on different GPUs, and conducted model validation and evaluation
- Designed efficient real-time decision algorithms to handle the impact of current decisions on future outcomes and optimized resource allocation

Research on Airport Congestion Contagion Shenzhen, China
Supervisor: Prof. Jianfeng Mao, School of Data Science, CUHKSZ Dec 2023 - Present

- Aimed to modify heterogeneous Susceptible-Infected-Susceptible (SIS) epidemic spreading to model airport congestion contagion
- Improved the epidemiological SIS model by transforming network transmission parameters into time-dependent functions and using the adaptive graph learning structure AdapGL model to learn the relevant parameters
- Introduced heterogeneous, dynamic, and negative recovery states into the SIS model and adopted adaptive graph learning methods to predict infection and recovery states
- Compared the improved SIS model with ODE methods, LSTM and ASTGCN models, showing that the adaptive graph learning approach (GAT+AdapGL) demonstrated superior performance in predicting airport delay statuses
- Proved that under specific conditions, delay propagation can still converge even when recovery and infection rates vary across different nodes
- Completed simulation experiments, including setting realistic infection rates and transmission processes, demonstrating the correctness of theoretical derivation

Optimization of Guangzhou Bus Rapid Transit (BRT) Public Road Usage Chapel Hill, NC
Supervisor: Prof. Guanting Chen, Dept. of Stat. & Operations Research, UNC Jan 2024 - Present

- Aimed to combine integer optimization and machine learning to solve the optimization problem of the Guangzhou BRT public road usage
- Applied LSTM, transformer, and random forest to fill in missing data
- Built an optimization model for the public road usage problem, using the active set method and null space method to determine the optimizable variables
- Explored plane-cutting and pruning techniques in integer optimization to accelerate the optimization solver
- Used the PyTorch framework to build and debug the model; compared the performance before and after applying the model to verify the effectiveness of combining integer optimization techniques with deep learning

Study on Flight Delay Propagation Modeling Shenzhen, China
Supervisor: Prof. Jianfeng Mao, School of Data Science, CUHKSZ Oct 2023 - Mar 2024

- Summarized the existing research methods on airport delay propagation and wrote a comprehensive review paper
- Sorted out more than 40 studies of statistics, econometric and queuing models on airport delay propagation and summarized the methods of each model category, the application of data scale, its own characteristics and the problems suitable for solving
- Wrote a paper titled "Flight Delay Propagation Modeling: Data, Methods, and Future Opportunities", which was accepted by the Transportation Research Part E (TRE)

Lithology Based Rock Prediction Shenzhen, China
Supervisor: Prof. Fangda Song, School of Data Science, CUHKSZ Mar 2023 - Jun 2023

- Aimed to predict the rock properties of undetectable places based on the variation of rock properties of detectable places

with depth
- Used regular expressions to extract three important features from a 14-dimensional rock dataset and employed Pandas to correlate these data to rock depth data, forming a new analyzable dataset
- Plotted depth curves to visualize the three features of different rock types, revealing that these three features can distinguish between different rock categories
- Trained rock data using SVM and HMM models; SVM exhibited high prediction accuracy but tended to overlook minority rock types, while HMM showed lower prediction accuracy but did not neglect minority rock types
- Planned to establish a new model with inputs including mine labels, rock features, and types, sampled by MCMC and optimized by the EM algorithm

COSCO Shipping Analysis Shenzhen, China
Supervisor: Prof. Jianhua Huang, School of Data Science, CUHKSZ Sept 2022 - Feb 2023
- Aimed to enhance the efficiency of maritime intelligent systems through regional division, route prediction, and congestion analysis
- Utilized the K-means model with spherical distance to cluster trade zones and divided the trade zones into nine regions based on the principle of minimizing errors, each corresponding to different trade categories
- Leveraged historical data of ships passing through ports to predict shipping routes using the textRNN model
- Conducted data cleaning and processing of three years' freight tonnage data at the Shanghai Port using Pandas and regular expressions
- Employed time series models ARMA with a moving window to forecast port congestion for the next month[See report]

COURSE PROJECT

Handwriting Recognition Shenzhen, China
Supervisor: Prof Xintong Li, School of Data Science, CUHKSZ Feb 2024 - Apr 2024
- Collected ten thousand images of human handwriting and cleaned image data by weakening the background color and horizontal lines
- Used OCR techniques to recognize word boundaries and automatically label these images, storing them in a specific structure for subsequent processing
- Applied transformer and LSTM architectures to learn the features of the images and evaluated the labels after recognition, improving the accuracy by 4%
- Modified the GAN recognition architecture to more accurately identify the fonts [see GitHub]

Study on the Algorithms of Singular Value Decomposition (SVD) Shenzhen, China
Supervisor: Prof Andre, School of Data Science, CUHKSZ Nov 2023 - Dec 2023
- Investigated and researched various algorithms for computing singular value decomposition (SVD) and applied the SVD algorithm to several interesting image applications
- Developed the SVD algorithm using two mathematical ideas without invoking Python software packages
- Created two types of fuzzy matrices to blur images, restored the images using my own SVD algorithms, and compared the restoration quality and runtime between different blur methods and different handwritten SVD algorithms

MODELING COMPETITION

Mathematical Contest In Modeling: Optimization of Great Lakes Water Level Shenzhen, China
Team Leader (received the Meritorious Winner award at the MCM) Jan 2024 - Feb 2024
- Collected historical data of more than ten factors such as evaporation and precipitation in the Great Lakes
- Applied Canonical Correlation Analysis (CCA) to obtain the impact matrix of environmental factors on water level changes, and combined DEMATEL and AISM algorithms to identify key elements and their impact levels
- Used the SARIMA model and Artificial Neural Networks (ANN) to capture and model the complex nonlinear changes, ultimately obtaining the dynamic network flow model of the Great Lakes
- Established water level-flow differential equations and added perturbation factors to improve model robustness
- Implemented Genetic Algorithm (GA) for multiple iterations to obtain the optimal control method
- Based on the static impact model and the dynamic network flow model, used the Sobol index method to analyze the sensitivity of the control algorithm to environmental changes[See report]

COMPUTER SKILLS

- **Programming:** Python (NumPy, Pandas, Matplotlib, Scikit-learn, Selenium), MATLAB, R
- **Software:** WIND, CSMAR, STATA, Excel
- **English Proficiency:** TOEFL 96