

Mr. Wan Zhuoyue

Birthday: 14/01/2000 | Mobile: +852 93692400 | E-mail: zhuoy.wan@connect.polyu.hk

Address: Building 3, Menglishuixiang Liuyuan Community, Nanchang County, Nanchang City, Jiangxi Province, China, 330200

Homepage: <https://zwanah.github.io/>

EDUCATION

Doctor of Philosophy in Computer Science, The Hong Kong Polytechnic University	01/2025-Now
Master of Science in Data-Driven Modeling, The Hong Kong University of Science and Technology	09/2022-08/2023
Bachelor of Science in Statistics, Chongqing University	09/2017-06/2021

RESEARCH EXPERIENCE

Research Assistant (Supervised by Asst. Prof. Chen Zhang, POLYU)	11/2023-11/2024
✓ Focused on developing pre-trained language models for handling Query Generation tasks;	
Trustworthy Medical Image Classification (Supervised by Asst. Prof. Hao Chen, HKUST)	10/2022-06/2023
✓ Providing rigorous and meticulous theoretical derivations of the method;	
✓ Conducting research as a Research Assistant;	
✓ Investigating topics such as noisy label, debias, fairness, and long-tail data in Medical Computer Vision;	
✓ Establishing a medical debias benchmark in mitigating bias in medical image classification;	
Medical Image Segmentation (Supervised by Prof. Yang Xiang, HKUST)	09/2022-06/2023
✓ Reproducing a Medical Image Segmentation paper that includes an Elastic Interaction-Based Loss Function;	
✓ Conducted multi-label segmentation in GI tract images for stomach, large bowel, and small bowel classes;	
✓ Extended a two-dimensional loss function to its three-dimensional form;	
Prediction of "High to turn" in the stock market (Supervised by Prof. Zhimin Zhang, CQU)	09/2020-06/2021
✓ Built theoretical models such as Stacked XGBoost, LightGBM, CatBoost, LR, RFC, and SVM models;	
✓ Conducted analysis of "High to turn" phenomenon and utilized predictive modeling to forecast its occurrence;	

PUBLICATIONS

DataVisT5: A Pre-trained Language Model for Jointly Understanding Text and Data Visualization	
✓ First Author, Accepted by ICDE 2025;	
✓ Available on ArXiv as a preprint (https://arxiv.org/abs/2408.07401);	
TransFlower: An Explainable Transformer-Based Model with Flow-to-Flow Attention for Commuting Flow Prediction	
✓ Second Author, Under review;	
✓ Available on ArXiv as a preprint (https://arxiv.org/abs/2402.15398);	
Medical Image Debiasing by Learning Adaptive Agreement from a Biased Council	
✓ Under review;	
✓ Available on ArXiv as a preprint (https://arxiv.org/abs/2401.11713);	

HONORS & AWARDS

Second Prize (National level), The Chinese Mathematics Competitions	11/2018
Third Prize (National level, Top 6%), The 8th TipDM Cup Data Mining Challenge Committee	06/2020
Second-class Scholarship (Top 2%), Chongqing University	05/2021
Third-class Scholarship (Top 5%), Chongqing University	11/2018
Advanced Individual of Scientific and Technological academic innovation (Top 1%), Chongqing University	12/2020
Outstanding Student (Top 1%), Chongqing University	01/2019

PROJECT EXPERIENCE

Predict the 2022 College Men's Basketball Tournament (Kaggle)	
✓ Predicted the 2022 College Men's Basketball Tournament using a logistic linear regression model;	
✓ Achieved excellent forecast results with a simpler model, showcasing strong analytical abilities;	
Team leader (3 people), the 8th "TipDM Cup" Big Data Mining Race(National-level)	07/2019-08/2019
✓ Awarded Third Prize at the national level for an analysis of the "High to turn" phenomenon in the stock market;	
✓ Led a team in building theoretical models such as BP neural network and Logistic models, as well as data collection and programming implementation using R;	
✓ Conducted in-depth analysis to predict the occurrence of "High to turn";	
Prediction of "High to turn" in the stock market based on Stacking Ensemble model	10/2020-06/2021
✓ Developed a Stacking Ensemble model using Python to predict the "High to turn" phenomenon in the stock market;	
✓ Improved classification performance by utilizing Stacked XGBoost, LightGBM, CatBoost, LR, RFC, and SVM;	

✓ Continued the previous project and achieved better results with the Stacking Ensemble model;

PROFESSIONAL SKILLS

Knowledge:	Strong mathematical background (Mathematical analysis, Advanced algebra, Numerical analysis, Partial differential equation, Real Analysis), data-driven background (Network modelling and Statistical machine learning) and optimized theory background (Information science, Operational research)
Programming:	Python, R, SPSS, SAS, Matlab, C++
Language:	IELTS: 6.5 (L: 7.0 R: 7.0 W: 6.5 S: 5.5)