Yufan Zheng

Email: ivan.zane1999@gmail.com or zhjpre@gmail.com | **Website**: https://yufanzheng.github.io
Scholar: https://scholar.google.com/citations?user=btkYKNAAAAAJ | **ORCID**: 0000-0003-0781-0308
Research Interests: Data Mining, Data Science, Machine Learning, AI for Science.

EDUCATION

NanFang College of Sun Yat-Sen University, Electrical and Computer Engineering Guangdong, China
Bachelor of Engineering Sep 2018 - Jun 2022

- **GPA**: 86.36/100. Major in Computer Science and Technology.
- **Scholarship**: Outstanding Graduates (2022), **Nation Scholarship** (2020).

PUBLICATIONS

[1] **Zheng Y**, Jiang W, Zhou A, Nguyen Q, Zhan C, Tong C. (2024). Epidemiology-informed Graph Neural Network for Heterogeneity-aware Epidemic Forecasting. arXiv preprint arXiv:2410.00049. (Under review in ICDE 2025)
[2] **Zheng Y**, Yue K, Wong, Eric W M, Yuan, H Y. (2024). Association between meteorological factors and human mobility with mosquito activity risk in Hong Kong. medRxiv. (Under review in Infectious Disease Modelling, JCR Q1)
[3] Zhan C, **Zheng Y**, Shao L, Chen G, Zhang H. (2023). Modeling the spread dynamics of multiple-variant coronavirus disease under public health interventions: A general framework. Information Sciences, 628, 469-487. (**JCR Q1**)
[4] Zhan C, Jiang W, **Zheng Y**, Lu J, Zhang Q. (2023). A data-driven study of active meteorological stations and the factors motivating their establishment. Sustainable Energy Technologies and Assessments, 57, 103147. (**JCR Q1**)
[5] Min H, Wu K, Tan M, Lin J, **Zheng Y**, Zhan C. (2022, July). Ensemble Learning for Crowdfunding Dynamics: JingDong Crowdfunding Projects. In International Conference on Neural Computing for Advanced Applications (pp. 372-386). Singapore: Springer Nature Singapore.
[6] Zhan C, **Zheng Y**, Zhang H, Wen Q. (2021). Random-forest-bagging broad learning system with applications for covid-19 pandemic. IEEE Internet of Things Journal, 8(21), 15906-15918. (**JCR Q1**)
[7] Zhan C, **Zheng Y**, Lai Z, Hao T, Li B. (2021). Identifying epidemic spreading dynamics of COVID-19 by pseudocoevolutionary simulated annealing optimizers. Neural Computing and Applications, 33, 4915-4928. (**JCR Q2**)
[8] **Zheng Y**, Zhen Q, Tan M, Hu H, Zhan C. (2021, November). COVID-19's Impact on the Box Office: Machine Learning and Difference-in-Difference. In 2021 16th International Conference on Intelligent Systems and Knowledge Engineering (ISKE) (pp. 458-463). IEEE.
[9] Wu S, Hu H, **Zheng Y**, Zhen Q, Zhang S, Zhan C. (2021, November). The Impact of COVID-19 on Online Games: Machine Learning and Difference-in-Difference. In CCF Conference on Computer Supported Cooperative Work and Social Computing (pp. 458-470). Singapore: Springer Nature Singapore.
[10] Li J, **Zheng Y**, Hu H, Lu J, Zhan C. (2021, November). Predicting Video Game Sales Based on Machine Learning and Hybrid Feature Selection Method. In 2021 16th International Conference on Intelligent Systems and Knowledge Engineering (ISKE) (pp. 497-502). IEEE.
[11] Lin J, Tan M, **Zheng Y**, Wu K, Zhan C. (2021, November). Detection Capability Prediction Based on Broad Learning System during the COVID-19 Pandemic. In 2021 16th International Conference on Intelligent Systems and Knowledge Engineering (ISKE) (pp. 697-702). IEEE.
[12] Wu S, **Zheng Y**, Lai Z, Wu F, Zhan C. (2019, October). Movie box office prediction based on ensemble learning. In 2019 IEEE Symposium on Product Compliance Engineering-Asia (ISPCE-CN) (pp. 1-4). IEEE.

RESEARCH EXPERIENCE

Independent Research: Graph Neural Network for Heterogeneous Epidemic Transmission. Jul 2024 – Present
Independent Researcher | Advisor: **Dr. Rocky Chen** (The University of Queensland)

- Characterized two heterogeneities of epidemics based on the mathematical model in the real-world dataset.
- Proposed a novel framework for heterogeneous epidemic forecasting that jointly discovers spatio-temporal and mechanistic heterogeneity and achieves state-of-the-art and effective performance across three real-world datasets.
- Developed more than ten prediction models as the baseline for epidemic forecasting, such as statistical time series models (AR, ARMA, GAR, etc.), deep learning models (DCRNN, LSTNet, CNNRNN-Res, etc.), graph neural network models (ST-GCN, TMGNN, Cola-GNN, EpiGNN, etc.).

City University of Hong Kong, Electrical and Engineering & Biomedical Sciences Department. Hong Kong, China
Full-time Research Assistant | Supervisor: **Dr. Eric Wong** Dec 2022 – Aug 2024
Responsibilities include assembling and curating datasets, building mathematical, statistical, and machine learning models, performing data statistical analysis and literature reviews, and writing publications, grants, and government reports.

1) Modeling Mosquito Activity Risks Based on Statistical Model.

Most mosquito-borne diseases are public health threats. To help monitor and prevent mosquito outbreaks, this research aimed to build prediction models to explore the association between different factors and mosquito activity.

- Developed the mosquito activity prediction models based on the statistical model (Distributed Lag Linear and Non-Linear Models) to explore the effect of human mobility and meteorological lag on mosquito activity prediction.
- Evaluated the model performance based on Step-WAIC and cross-validation and implemented sensitivity analysis for model parameters.

Collaborative Research: Intervention Evaluation based on Computational Epidemiology. Aug 2021 – Aug 2024
Independent Researcher | Advisor: **Dr. Choujun Zhan** and **Prof. Guanrong Chen** (City University of Hong Kong)

1) Epidemic Transmission Modeling based on Complex Networks and Dynamic Systems.

- Proposed a network-based epidemiological model (system dynamics model), which considered multiple antibodies of individuals and partial cross-immunity between different viruses, and then was simulated in different theory networks (random networks, scale-free networks, small-world networks, etc.).

2) Epidemic Intervention Evaluation based on Dynamic Systems.

- Proposed an epidemiological framework (based on the dynamic systems model) for simulating the multi-directional mutation process and transmission under the scenario considering multiple variants and massive vaccinations.
- Evaluated single and combined public health interventions, including non-pharmaceutical, pharmaceutical, and vaccine interventions, based on sensitivity analysis.

NanFang College of Sun Yat-Sen University, Electrical and Computer Engineering Department. Guangdong, China
Research Assistant | Supervisor: **Dr. Choujun Zhan**, Advisor: **Prof. Haijun Zhang** Mar 2019 - Mar 2022

1) Modeling in Healthcare and Public Health based on Machine Learning and Dynamic Systems.

- Constructed and cleaned infectious disease time-series data containing 184 countries and 1241 regions from December 2019 to October 2021, obtained from public health departments and multiple data sources.
- Applied statistical models to quantify the impact of the infectious disease outbreak on the box office and online game players, and machine learning models for their prediction.
- Improved an epidemiological model (dynamic systems model) combining intercity migration networks to describe the intercity transmission and developed an improved simulated annealing method to optimize the model parameters.
- Proposed a novel model combining ensemble learning and broad learning systems to predict epidemic prediction.
- Developed machine learning models incorporating explanation (Shapley Value) in healthcare resource prediction.

2) Modeling in Entertainment based on Machine Learning.

- Crawled and constructed historical video game sales data from 1970 to 2018, a global movie statistics platform from 1980 to 2017, and Chinese Box office time-series data from 2011 to 2019 based on Web crawlers.
- Applied data manipulation, cleaning, visualization, and correlation analysis.
- Proposed a novel hybrid feature selection method to improve the performance of video game sales prediction.
- Developed the box office prediction model based on ensemble learning.

WORK EXPERIENCE

Huangpu Institute of Materials, Industrial Software Development Division Guangdong, China
Technician Mar 2022 - Nov 2022

Responsibilities include data processing and feature engineering, algorithm model development and optimization, algorithm effect evaluation, technology research and innovation, and cross-departmental collaboration.

- Designed an industrial drawing recommendation framework, including collaborative filtering based on a drawing similarity matrix and a method based on cluster and classification models.
- Developed blood pressure monitoring model and road condition detection model based on machine learning with sensor data and participated in the research of the stocker controller system requirements.

AWARDS AND PROJECTS

Awards

- Provincial third prize, awarded by China Undergraduate Mathematical Contest in Modeling. Oct 2021
- Merit award paper in 2019 IEEE International Symposium on Product Compliance Engineering-Asia. Oct 2019

Funded Projects

- Co-PI, Provincial College Students' Innovative Entrepreneurial Training Plan Program, China (10,000 CNY). 2021
- PI, Provincial College Students' Innovative Entrepreneurial Training Plan Program, China (10,000 CNY). 2019

ACADEMIC ACTIVITIES

Conference experience

- Best volunteer at the International Conference on Neural Computing for Advanced Applications 2021, held in Guangzhou, China, led the volunteers' group in preparation and implementation. Aug 2021
- Participated in and delivered an oral presentation at the 2019 IEEE International Symposium on Product Compliance Engineering-Asia in Hong Kong, China. Oct 2019

Review experience

- Internet of Things.
- Infectious Disease Modelling.

ADDITIONAL INFORMATION

References

- Dr. Choujun Zhan, South China Normal University, Associate Professor: zchoujun2@gmail.com.
- Prof. Haijun Zhan, Harbin Institute of Technology, Shenzhen, Professor: hjzhang@hit.edu.cn.

Programming

Python (PyTorch, Tensorflow, Scikit-Learn, Numpy, Pandas), R, MATLAB, C, Java, MySQL, Linux, Git, LaTeX.