PUYANG ZHAO

(+1) 8585006534 ⋄ puyang.zhao@uth.tmc.edu⋄ https://puyangzhao.github.io

EDUCATION

Ph.D. Candidate in Biostatistics and Data Science *Aug 2023 - Present*
The University of Texas Health Science Center at Houston
Expected to receive Ph.D. in Biostatistics and Data Science.

M.S. in Mathematics and Statistics *July 2019 - July 2021*
The University of Melbourne
Specialization in Statistics and Stochastic Processes

B.S. (Honours) in Statistics *September 2015 - June 2019*
Beijing Normal University and Hong Kong Baptist University United International College (BNU-HKBU UIC)
Received B.S. (Honours) in Statistics from Hong Kong Baptist University

PUBLICATIONS

Puyang Zhao, James J. Yang, Anne Buu* "Applied statistical methods for identifying features of heart rate that are associated with nicotine vaping" has undergone a minor review and been returned to The American Journal of Drug and Alcohol Abuse.

Huiyun Zhang*, Heming Huang, **Puyang Zhao**, Xiaojun Zhu, Zhenbao Yu "CENN: Capsule-Enhanced Neural Network with Innovative Metrics for Robust Speech Emotion Recognition" has been published in Journal of Knowledge-Based Systems September 2024.

Puyang Zhao, Xinhui Liu, Zhiyi Yue, Qianyu Zhao, Xinzhi Liu, Yuhui Deng, Jingjin Wu*. "DiGAN Breakthrough: Advancing Diabetic Data Analysis with Innovative GAN-Based Imbalance Correction Techniques" has been published in Journal of Computer Methods and Programs in Biomedicine Update, 2024.

Puyang Zhao, Wei Tian, Lefu Xiao, Xinhui Liu, Jingjin Wu. "An Attention-based Long Short-Term Memory Framework for Detection of Bitcoin Scams" Accepted by IEEE International Conference on High Performance Big Data and Intelligent Systems 2022 (Best Paper Nomination Award)

Puyang Zhao and Wei Tian. "Research on Prediction of Solar Power Considering the Methods of Statistical and Machine Learning–Based on the Data of Australian Solar Power Market." IOP Conference Series: Earth and Environmental Science. Vol. 1046. No. 1. IOP Publishing, 2022.

TEACHING EXPERIENCE

Graduate Teaching Assistant September 2024 - December 2024
The University of Texas Health Science Center at Houston (UTHealth)

- Assisted in teaching and tutoring for PH 1700 Intermediate Biostatistics
- Graded homework assignments, unit tests, and final examinations
- Provided additional support to students during office hours

Assistant Instructor I *September 2021 - December 2021*
BNU-HKBU United International College

- Participated in teaching and tutoring for five subjects: Calculus I (1002), Calculus I (1004), Logistics, Network and Transportation Models, Data Analysis for Business (1001).
- Prepared lesson plans, follow-up exercises, homework assignments, unit tests, and final examinations.

PRESENTATIONS

Puyang Zhao, Wei Tian, Lefu Xiao, Xinhui Liu, Jingjin Wu. "An Attention-based Long Short-Term Memory Framework for Detection of Bitcoin Scams." Presented as the first author and nominated for the Best Paper Award at the IEEE International Conference on High Performance Big Data and Intelligent Systems 2022.

RESEARCH EXPERIENCE

Graduate Research Assistant *August 2023 - August 2024*
The University of Texas Health Science Center at Houston

- Conducted research on the association between biobehavioral indicators and vaping behavior. Collected and analyzed biobehavioral data, such as heart rate, during e-cigarette use. Applied signal processing and time series analysis techniques to extract features and patterns from the data. Investigated the potential and accuracy of biobehavioral indicators in detecting and predicting vaping behavior.
- Processed and cleaned experimental data to ensure data quality and utilized statistical methods for data analysis and modeling. Wrote research reports and academic papers detailing research methods, findings, and conclusions.

Research Assistant *February 2022 - August 2022*
Baylor University

- Conducted research on the performance of deep learning on imbalanced data and the difference of information entropy for imbalanced data.
- Utilized LSTM for predictive analysis of biological data (water temperature and dissolved oxygen prediction).
- Performed visualization and dimensionality reduction of gravitational wave data.
- Vectorized and analyzed financial text data.

Research Assistant *July 2021 - December 2021*
BNU-HKBU UIC

- Developed a new machine learning method 'A-LSTM', which achieved an F1 score of over 82% on the original data, surpassing existing methods.
- Proposed a novel data mining method 'DiGAN', combining Generative Adversarial Network (GAN) and Random Forest (RF), to classify a public dataset with 5,070 records. DiGAN achieved an F1-score up to 98%, representing at least 4% improvement over traditional techniques.

Research Assistant *December 2019 - February 2020*
North Carolina State University

- Implemented R code to draw samples from Forward Filtering-Backward sampling and compared the results with the marginal probability calculated by the Forward-Backward Algorithm method.
- Applied Forward Algorithm in Rcpp and compared the results with those from the Forward Algorithm implemented in R code. Completed the proof of Forward Filtering-Backward sampling.

Master's Thesis *December 2019 - May 2021*
University of Melbourne

- Complemented the relevant solar energy data through various solar irradiance models.
- Proposed an effective method for accurate parameter estimation of photovoltaic (PV) array models by using Long Short-Term Memory (LSTM) and a time series model (a multiple regression model with correlated errors) to analyze Ausgrid's Solar Home Electricity Data from July 1, 2010, to June 30, 2013.
- Demonstrated that these techniques performed well in prediction compared to traditional time series models.

Undergraduate Thesis *June 2018 - December 2018*
BNU-HKBU UIC

- Applied the ARIMA model, Holt-Winters model, and the prophet package to forecast mobile traffic based on real data. Proposed a base station sleeping strategy to reduce power consumption based on the prediction.
- Utilized Monte-Carlo simulation to seek an optimal base-station sleeping strategy to maximize energy efficiency of the network.
- Developed a base-station sleeping model based on switch control of the base stations, discharging data to overlapping stations when data requirements were satisfied to minimize power consumption.
- Applied power optimization algorithms to control and optimize the switch of stations. Simulation results on Matlab showed optimized station utilization and power utilization after processing the algorithm.

LEADERSHIP ACTIVITIES & COMMUNITY SERVICE

Student Affairs Office *September 2017 - June 2019*

- Spearheaded the planning and execution of diverse campus events, including formal high table dinners and engaging dormitory activities
- Managed and enhanced the WeChat Public Account for the Scientific Education Hall, driving increased student engagement and awareness
- Demonstrated strong organizational and communication skills through successful event coordination and digital media management

United Innovation Charity Club *September 2015 - June 2016*

- Engaged in the Mangrove Forest program for outdoor activities and voluntary teaching, with 35 volunteer hours.
- Liaised with external organizations for sponsorship and activity funding.

Nepal International Voluntary Teaching Project *January 2016*

- Participated in teaching and tutoring the Chinese language.
- Prepared lesson plans, follow-up exercises, homework assignments, and unit tests.

SKILLS & LANGUAGES

Software: Python, R, SAS, STATA, GIS, Julia, Matlab, LaTeX, C++, MySQL, Rcpp, Microsoft Office Specialist
Languages: Mandarin (Native), English (Fluent)

HONORS & AWARDS

AMSI Summer School in the Mathematical Sciences *January 2021 - February 2021*
Joined the course "The Mathematical Engineering of Deep Learning" with First Class Honours.

C++ Programming for Financial Engineering with Distinction *February 2022 - June 2022*
Offered by the Baruch College's Financial Engineering (MFE) program as part of the Pre-MFE Program.

National Mathematical Modeling Competition for College Students *September 2018*
Award: Third Prize

STATISTICS COURSES

Advanced Statistics, Applied Stochastic Process, Categorical Data Analysis, Computational Statistics and Data Science, Data Analysis Using R, Mathematical Statistics, Multivariate Analysis, Multivariate Statistical Techniques, Optimization, Random Processes, Regression Analysis, Simulation, Statistical Modeling, Stochastic Modeling, Structured Programming, Survey Sampling, Time Series Analysis, Statistical Inference