Zhongyuan Zhao

Email: zzhao62@binghamton.edu / Tel: 607-766-8859

Education

Binghamton University(SUNY), Department of Mathematics and Statistics — **Binghamton, NY**
Ph.D. Candidate in Mathematics and Statistics — Aug 2021 – Present

Binghamton University(SUNY), Department of Mathematics and Statistics — **Binghamton, NY**
Master of Arts in Mathematics — Aug 2021 – May 2023

Jilin University, School of Mathematics — **Jilin, China**
Bachelor of Science in Mathematics and Applied Mathematics — Sep 2016 – Jun 2020

Technical Skills

Programming: Python, R, MATLAB

ML/DL Frameworks: PyTorch, scikit-learn, TensorFlow, Hugging Face Transformers

Publication & Poster Presentation Publication

1. B. Yi, S. Chen, M. A. Rahman, S. H. Bea, M. Bashar, Z. Zhao, T. Zhang, "A PBPK Model for Prediction of Apixaban and Rivaroxaban Secretion into Milk" (Submitted)
2. Z. Zhao, P. Zou, Y. Fang, T. Si, B. Yi, T. Zhang, “Machine Learning Approaches for Assessing Medication Transfer to Human Breast Milk: Estimating Milk/Plasma Drug Ratios” (Under review: Journal of Pharmacokinetics and Pharmacodynamics)
3. A. Stratigakis, D. Paty, P. Zou, Z. Zhao, Y. Li, and T. Zhang, “A regression approach for assessing large molecular drug concentration in breast milk,” Reproduction and Breeding, vol. 3, no. 4, pp. 199–207, 2023.

Poster Presentation

1. “Machine Learning Approaches for Assessing Medication Transfer to Human Breast Milk: Estimating Milk/Plasma Drug Ratios”, AAPS 2024 PharmSci 360, Salt Lake City, UT, Oct 2024

Research Experience

Fake News Detection Using BERT (In Progress) — 07/2025 - Now

- Currently fine-tuning a BERT model (bert-base-uncased) on around 40k news headlines
- Task: binary classification (FAKE vs REAL) using Hugging Face Trainer API
- Focused on optimizing F1-score and integrating model explainability
- **Tools Used: Python, Hugging Face Transformers, PyTorch**

Predicting Human Bioavailability of Monoclonal Antibodies from Antibody Sequences — 05/2025 - Now

- Planning to develop machine learning models (CNN, Transformer-based methods) to predict human bioavailability and pharmacokinetic parameters of subcutaneously administered monoclonal antibodies directly from antibody sequences
- Aim to computationally extract sequence-derived physicochemical and biophysical features (isoelectric point, hydrophobicity, aggregation propensity) for predictive modeling
- Will integrate non-human primate pharmacokinetic data to improve prediction accuracy
- **Tools to be used: Python, PyTorch, Biopython**

Research On Qusi-Stationarity of the Shiryaev Recurrence in Exponential Case (Ph.D. Dissertation) — 05/2023 - Now

- Find tight bounds for the quasi-stationary function and ADD(Average Delay of Detection) in Shiryaev Robert change point detection method which corresponds to the accuracy of change point detection
- Find the boundary for the control parameter in Shiryaev Robert change point detection
- Prove the optimality of the Shiryaev-Robert Procedure
- Do simulation of quasi-stationary function, ADD, and control parameter
- **Tools Used: Matlab, Python**

- **Paper under prepare**

Build a PBPK Model for Prediction of Apixaban and Rivaroxaban Secretion into Milk 05/2023 - 10/2024

- Led statistical modeling and data analysis for a whole-body PBPK lactation model to predict drug secretion of apixaban and rivaroxaban into breast milk
- Validated model predictions using clinical concentration–time profiles and calculated accuracy metrics (Tmax, Cmax, AUC, M/P ratio), achieving <10% prediction error
- **Tools Used: R**
- **Paper [1] Submitted**

ML Model to Predict Drug Secretion Ratios (M/P) using Clinical Features 08/2023 - 08/2024

- Using PCA to visualize the data focus on MP ratio distribution(Milk/Plasma) behavior
- Build KNN(K-Nearest Neighborhood), SVM(Support Vector Machine), RF(Random Forest), and NN(Neural Network) models to do classification for MP ratio
- Using a grid search combined with cross-validation to find the best tuning parameters for each method, combine with different attempts like normalizing the data, improve the average accuracy by around 20%
- Visualize machine learning models and compare by accuracy
- **Tools Used: Python**
- **Paper [2] under review**

Multivariate Regression for Large Molecule Drug Modeling 05/2023 - 10/2023

- Developed a multivariate regression model to predict milk-to-plasma (M/P) drug concentration ratios of monoclonal antibodies (mAbs) using physicochemical and pharmacokinetic properties
- Conducted data curation from DrugBank, PubMed, and Antibody Society databases for 11 mAbs with clinical milk concentration data
- Applied the final model to predict M/P ratios for 79 approved therapeutic antibodies, supporting risk assessment of biologic drugs during lactation
- **Tools Used: R**
- **Paper [3] published**

Implementation of improved Fused Lasso Image Denoising 08/2022 - 12/2022

- Implemented the 2D Fused Lasso algorithm for grayscale image denoising via coordinate descent and group fusion strategy
- Designed and optimized a custom objective function including MSE, L1, and total variation penalties
- Simulated and tested on various noise patterns (Gaussian, uniform, structured) to evaluate denoising performance
- **Tool Used: R**

Work Experience

Research Assistant, Binghamton University - Binghamton, NY 05/2023 - Now

- Research on machine learning and AI algorithms combined with clinical data from breast milk to predict PK parameters
- Building PBPK model through **Python** and **R**
- Data visualization and analysis to support other projects

Calculus Instructor & Statistics TA, Binghamton University - Binghamton, NY 08/2023 - Now

- Instructor of calculus(Math 224, 225, 226, 227) and teaching assistant of Elementary Statistics(Math 147)
- Holding office hours

Research Assistant, Nankai University - Tianjin, China 06/2019 - 08/2019

- Research on Low-level Vision denoising Problems using Dual Convolutional Neural Networks
- Using **MATLAB** to apply Dual Convolutional Neural Networks to do image deraining