# YICHI ZHANG

Google Scholar ⋄ Github

Phone: (+86) 19556575257 ⋄ Email: yichi.zhang@stonybrook.edu

## EDUCATION

**State University of New York at Stony Brook (SBU)** *August 2024-May 2026(expected)*
B.S. in Applied Math and Statistics
GPA: 3.92/4.0
*Related courses: Data Mining, Data Analysis, Stochastic Model, Complex Analysis.*

**Anhui University (AHU)** *September 2022-July 2024*
B.S. in Applied Statistics
GPA: 3.79/5.0 (87.9/100)
*Related courses: Probability Theory and Mathematical Statistics, Ordinary Differential Equation (ODE), Abstract Algebra, Object-Oriented Programming (Java), Data Structure (Java).*

## RESEARCH INTERESTS

I am interested in generally AI4Science. Focus on Generative Model and LLM.

## RESEARCH EXPERIENCE

**RL-Guider: Leveraging Historical Decisions and Feedback for Drug Editing with LLMs**
*Supervisor: Prof. Yi Liu* SBU

· Proposed RL-Guider, a reinforcement learning agent that guides large language models (LLMs) by leveraging rich historical feedback to enhance drug-editing tasks.
· Designed the reinforcement learning framework to provide strategic guidance, effectively improving the performance of LLM-based molecular editing.
· Demonstrated significant improvements in decision-making quality through systematic experiments and evaluations.

## PUBLICATIONS

Xufeng Liu, Yixuan Ding, Jingxiang Qu, **Yichi Zhang**, Wenhan Gao, and Yi Liu, "RL-Guider: Leveraging Historical Decisions and Feedback for Drug Editing with Large Language Models", *ACL, under review*, 2025.

## ACHIEVEMENTS

| | |
|---|---:|
| Scholarship, awarded by AHU | *Fall 2023* |
| MCM/ICM Honorable Mention , awarded by COMAP | *Spring 2024* |

## SKILLS/HOBBIES

| | |
|---|---|
| **Programming Languages** | Python, Java, Latex |
| **Framework & Library** | Pytorch, Pandas, Numpy, Matplotlib, RDKit |
| **Language** | Mandarin and English (TOEFL 102, GRE 320) |

## PROJECTS

### CNN Classification on CIFAR-10 Using VGG16 · Jan 2025

- Fine-tuned pretrained VGG16 model in PyTorch for CIFAR-10 image classification.
- Utilized data augmentation and learning rate scheduling to enhance model generalization and achieve 93% accuracy.
- Visualized training progress and performance metrics using Matplotlib.

### Sentiment Analysis Using LSTM · Jan 2025

- Developed and trained a two-layer LSTM model for sentiment classification tasks.
- Implemented data preprocessing pipelines to optimize dataset quality.
- Monitored model training with tqdm for efficiency and accuracy improvement, and ultimately achieve 90% accuracy.

### Generative Modeling with VAE · Feb 2025

- Built and optimized a Variational Autoencoder (VAE) for digit generation on MNIST.
- Leveraged ELBO loss function combining MSE and KL-divergence.
- Visualized latent spaces and generated samples for performance assessment.

### Illegal Wildlife Trade Data-Driven Project (MCM/ICM Paper) · Jan 2023

*Supervisor: Prof. Ranchao Wu* · AHU

- Developed a data-driven project to reduce illegal wildlife trade by selecting UNEP as the optimal client using AHP and TOPSIS.
- Proposed a five-year GGIWT program covering monitoring, law enforcement, education, policy, and international cooperation.
- Built a Multiple Linear Regression Model to predict project impact, demonstrating improved outcomes with low sensitivity and high accuracy.
- Applied data analysis techniques such as PCA and predictive modeling to support decision-making.