Ziyu Li

ziyul6@illinois.edu — +1(217)318-6593

Personal Statement

I am an undergraduate student at the University of Illinois Urbana-Champaign, majoring in Mathematics with a minor in Computer Science. I am interested in leveraging novel machine learning and artificial intelligence strategies to advance synthetic biology and protein engineering.

Educational Background

University of Illinois Urbana-Champaign
Bachelor of Science in Mathematics - *Data Optimization Concentration*
Minor in Computer Science *Jan 2024 – June 2026 (expected)*
- GPA: 4.0/4.0

Tongji University
Mechanical Engineering - *Mechatronics Concentration* *Sep 2021 – Jan 2024*

Experiences

Zhao Group @ Carl R. Woese Institute for Genomic Biology *Champaign, IL*
Undergraduate Researcher *Sep 2024 – Present*
- Constructing biological knowledge graphs, integrating biological datasets to support subsequent machine learning tasks
- Exploring representations of protein dynamics to enhance the effectiveness of graph representations of protein structures
- Developing multi-modal large concept model to explore the entire protein 'universe'

Illinois Data Science Club *Champaign, IL*
Education Lead *May 2024 – Present*
- Developing educational resources and staying updated on data science trends
- Mentoring club members and teaching them data science
- Organizing workshops and drove improvements in the educational experience

2025 Undergraduate Research Symposium (URS) *Champaign, IL*
Presenter *April 24, 2025*
- Presented my research, "Graph Neural Network for Enhanced Protein Solubility Prediction."

Projects

A Multi-Stage Learning Framework for Credit Card Spending Prediction, Risk Segmentation, and Limit Recommendation *Champaign, IL*
Modeler *Mar 2025*
- Built time series models with rich feature engineering for spending prediction
- Applied clustering (DBSCAN, BIRCH, Hierarchical) with neural network autoencoder-based dimensionality reduction
- Designed a hybrid rule-based & Random Forest recommendation system for credit limit adjustment
- Delivered actionable insights for low-risk growth and risk control

Interpretable Spatiotemporal Neural Networks for EEG-Based Bilingual Proficiency Prediction *Champaign, IL*
Modeler *Nov 2024*
- Analyzed EEG(electroencephalography) signals, identifying peak activity in bilinguals for efficient language processing
- Developed a CNN-LSTM model using EEG data to predict bilingual proficiency, capturing spatial and temporal brain activity patterns
- Used Integrated Gradients to highlight key brain regions (left frontal, right parietal, and right posterior temporal lobes) in relevant language processing tasks

Car Sales Recommendation System *Champaign, IL*
Lead Data Analyst *Feb 2024 – May 2024*
- Cleaned, processed, and merged demographic and car data
- Applied Bayesian ridge, Random Forest, gradient boosting classifier, and elastic net to predict car features based on demographics
- Constructed a user system for car dealerships to use this system more efficiently

Non-destructive and Efficient Prediction of Intramuscular Fat in Live Pigs Based on Ultrasound Images and Machine Learning *Nanjing, China*
Undergraduate Researcher *Aug 2023 – Dec 2023*
- Crafted predictive models for pig IMF based on ultrasonic images
- Applied swift, non-invasive methods to aid genetic improvement in pork
- Contributed to paper drafting and data analysis

Impact of Maternal Physical and Mental Health on Infant Development: A Machine Learning and Optimization Model Approach *Shanghai, China*
Researcher *Aug 2023*
- Leveraged Random Forest, SVM, and SVR algorithms to analyze and predict relationships between maternal health indicators and infant development outcomes, including behavior and sleep quality
- Developed an Integer Programming model and applied Entropy Weight and Rank Sum Ratio (RSR) Evaluation to optimize treatment costs and classify infant sleep quality based on multiple health indicators
- Applied Ordinal Priority Approach (OPA) to prioritize sleep intervention methods, assessing each by safety, sustainability, and effectiveness

Innovative Development of an IRB Robotic Arm Replica and a Bluetooth Control Model *Shanghai, China*
Researcher *Feb 2023 - May 2023*
- Replicated a robotic arm's components using SolidWorks and a 3D printer
- Assembled the robotic arm with components and standard parts
- Developed a Bluetooth control system for remote use

Oil Temperature and Pressure Control System *Shanghai, China*
Researcher *Nov 2021 - Jan 2022*
- Developed an automated control system using LabVIEW to regulate oil temperature and pressure in mechanical systems
- Implemented real-time monitoring and dynamic adjustments based on sensor data for optimal system performance
- Enhanced system reliability and efficiency by automating feedback loops and ensuring precise control within safety limits

Honors & Awards

1st place (1/180) in 2025 Illinois Datathon *Champaign, IL*
Contestant *Apr 2025*
- Built time series models with rich feature engineering for spending prediction
- Applied clustering (DBSCAN, BIRCH, Hierarchical) with neural network autoencoder-based dimensionality reduction
- Designed a hybrid rule-based & Random Forest recommendation system for credit limit adjustment
- Delivered actionable insights for low-risk growth and risk control

LAS Dean's list (Fall 2024) *Champaign, IL*
Top student *Jan 2025*

LAS James Scholar *Champaign, IL*
Honor Program *Aug 2024 - Present*
- Academic excellence and a commitment to engaged inquisitive learning
- Global citizenship, sustainability, and interdisciplinary growth

2nd place (2/51) in 2024 Sandia National Lab Data Challenge *Champaign, IL*
Contestant *Nov 2024*
- Developed a CNN-LSTM model to predict bilingual proficiency from EEG signals, identifying key brain regions (left frontal, right parietal, and right posterior temporal lobes) involved in language processing and using Integrated Gradients to highlight spatial and temporal activity patterns for efficient interpretation. The final model has over 94% accuracy.

LAS Dean's list (Spring 2024) *Champaign, IL*
Top student *June 2024*

National First Prize in Mathematical Modeling *Shanghai, China*
Team Leader & Modeler *Aug 2023*
- Led a team to develop machine learning models assessing maternal health impact on infant development

Tongji University Undergraduate Merit-based Scholarship *Shanghai, China*
Top Student *Nov 2022*

37th Chinese Physics Olympiad (CPhO): First Prize in Jiangsu province, Second Prize nationally *Nanjing, China*
Contestant *Nov 2020*

Skills

- **Languages:** English (Fluent), Mandarin Chinese (Native)
- **Computer:** Python, R, Java, C++, MATLAB, LabVIEW, Arduino, SolidWorks