Xiangyu Zhang

✉ xzhan344@jh.edu ⌂ github.com/Papillon-Xiang

EDUCATION

Johns Hopkins University · Baltimore, MD
Master of Science in Engineering in Data Science; Entranced Scholarship · *Aug. 2023 – May. 2025(Expected)*
Core course: Bayesian Analysis, Machine Learning: Deep learning, Natural Language Processing, Machine learning for Medical Application, Computational Protein Structure Prediction and Design

Xi'an Jiaotong-Liverpool University · Suzhou, China
Bachelor of Science in Applied Mathematics; GPA: 3.82/4.00 · *Sep. 2019 – Jul. 2023*
Core course: Statistics and Probability, Applied Probability, Partial Differential Equation, Statistical Distribution Theory, Operational Research, Numerical Analysis, Metric Space, Mathematical Analysis

Imperial College London Data Science Winter School
Best NLP team for Imperial Data Science Winter School · *Jun. 2022 – Feb. 2022*

RESEARCH INTEREST

My research is motivated by understanding and deriving useful information from data in the biomedical field, through the lens of statistics and artificial intelligence.I am particularly interested in advancing decision-making in precision medicine and exploring generative AI techniques for drug discovery.

CONFERENCE

- **Xiangyu Zhang** (2024). Enhancing Molecular Design through Graph-based Topological Reinforcement Learning. *MOML*, 2024.

- **Xiangyu Zhang**, Brian Caffo*, Mark Shelhamer* (2024). Integrative Analysis in Spaceflight Physiology. *NASA's AWG Meeting*, 2024.

PUBLICATION

- Xinyu Chang, **Xiangyu Zhang**, Haoruo Zhang, Yulu Ran (2024). Music Emotion Prediction Using Recurrent Neural Networks. *arXiv Preprint*: https://arxiv.org/abs/2405.06747.

- **Xiangyu Zhang**, Yanxun Xu* (2024). Reinforcement Learning for Treatments of Sepsis Patients. (In preparation)

- Guannan He, **Xiangyu Zhang**, Xinyu Chang (2024). MambaDiff: Revolutionizing Seq2Seq Models with Diffusion Model and Mamba Architectures. *Poster link*. (To be submitted)

RESEARCH EXPERIENCE

Reinforcement Learning for Treatments of Sepsis Patients · Baltimore, Maryland
Researcher · *May. 2024 - Present*
Supervisor: Professor Yanxun Xu, Department of Applied Mathematics and Statistics, JHU

- **Temporal Dependency Modeling:** Proposed a GPT decoder-only architecture with causal masking to improve temporal modeling, leveraging self-attention to capture long-range dependencies and overcome the limitations of traditional Markov decision processes.
- **Clinical Interpretability:** Enhanced model interpretability of decision-making by incorporating clinician guidelines into Q-learning.

CIPHER(NASA): Integrative Analysis in Spaceflight Physiology · Baltimore, Maryland
Researcher · *Feb. 2024 - Present*
Supervisor: Professor Brian Caffo, Department of Biostatistics, JHU; Professor Mark J. Shelhamer, Department of Medicine, JHU

- **Inter-System Relationships**: Conducted Bayesian Canonical Correlation Analysis (VBCCA) to explore the changes in physiological relationships across ocular, cardiovascular, and skeletal systems prior and after spaceflight
- **MCMC sampling**: Performed pairwise analysis to examine correlations and variances, complementing results with MCMC-based prediction intervals

Molecular Optimization using Reinforcement Learning
Independent Researcher
 Accepted at the 5th Molecular Machine Learning Conference

Baltimore, Maryland
Aug. 2024 - Nov. 2024

- **MDP for Molecular Design** Formulated molecular modification as a Markov Decision Process (MDP) by implementing targeted atomic and molecular bond changes on SMILES representations.
- **State Construction:** Designed a state space integrating multi-scaled colored graphs, persistent homology, and chemical descriptors using molecular fingerprints.
- Designed the reward function with specific properties while maintaining similarity to the original molecule.

Enhancing Seq2Seq with Diffusion Models and Mamba Architecture
Researcher
Supervisor: Professor Daniel Khashabi, Department of Computer Science, JHU

Baltimore, Maryland
Mar 2024 - May 2024

- **Model Integration**: Enhanced generation capabilities and optimized performance by integrating DiffuSeq models with Mamba's fast, adaptable architecture, transitioning the diffusion model of DiffuSeq from Transformer to Mamba
- **Seq2Seq Task Applications**: Preparing to evaluate our model across four major seq2seq tasks: open-domain dialogue, question generation, text simplification, and paraphrasing, aiming to validate versatility and effectiveness.

Music Emotion Prediction Using Recurrent Neural Networks
Researcher
Supervisor: Professor Anthony Kearsley, Department of Applied Mathematics and Statistics, JHU

Baltimore, Maryland
Feb 2024 - May 2024

- **Emotion Categorization**: Applied Russell's Emotion Quadrant to categorize music into four distinct emotional regions and developed models for accurate prediction of these categories. Enhanced model robustness by augmenting data through noise injection, time shifting, and pitch modification.
- **Feature Extraction**: Employed Librosa to compute the chromagram of each audio clip, converting augmented 3600 30-second audio clips into 14 distinct features for analysis.
- **Model Training and Evaluation**: Implemented and evaluated models using Recurrent Neural Networks (RNN), bidirectional RNNs, and LSTM, incorporating cross-validation techniques to enhance predictive accuracy.

Summer Undergraduate Research Fellow
Researcher
Supervisor: Professor Fangting Zheng, Department of Mathematical Science, XJTLU

Suzhou, China
June 2022 - Sep 2022

- **GNNs:** Investigated GNNs architecture model for analyzing citation network between Computer Science arXiv papers.
- **TDA:** Applied Topological Data Analysis to predict the protein-protein interactions (PPIs)

PROFESSIONAL EXPERIENCE

Knot Theory and Application Seminar
Member

Dalian, China
Jul 2023

- Delved into fundamental and advanced aspects of topology, focusing specifically on the applications of Topological Data Analysis (TDA) within the biomedical field.

PROGRAMMING SKILLS

Machine Learning: PyTorch, scikit-learn

Programming: Python, PostgreSQL, R, MATLAB, Java