Haoyu Li

330 De Neve Dr | (424) 293-9235 | haoyuli02@ucla.edu

Education

University of California, Los Angeles (UCLA) Sep 2020 - Jun 2024
Mathematics of Computation (Expected)

- **GPA**: 3.975/4.00 | Dean's Honor List (All quarters)
- **Relevant Courses:** Machine Learning, Optimization, Algorithms, Honors Numerical Analysis, Probability Theory, Honors Algebra, Honors Analysis, Software Development

Research Interests

- Machine Learning, Explainable AI, Generative AI, Graph Neural Network, AI4Science

Publications

- [1] **Haoyu Li***, Shichang Zhang*, Longwen Tang, Matheiu Bauchy, Yizhou Sun, "Predicting and Interpreting Energy Barriers of Metallic Glasses with Graph Neural Networks", Accepted at Neurips 2023 AI4Mat Workshop
- [2] Tong Xie*, **Haoyu Li***, Andrew Bai, Cho-Jui Hsieh, "Interpretability through Training Samples: Data Attribution for Diffusion Models", Under Review at CVPR 2024
- [3] Haoran Jia, **Haoyu Li**, Xinyue Li, Xiaoxian Shen, Yichen Wang, Zichun Liao, Andrea L. Bertozzi, P. Jeffery Brantingham, Jona Lelmi, "Intentional Youth Development Activities and Peer Effects in a Gang Prevention Program", Accepted at IEEE 2023 Big Data DS4EIW workshop

Research Experiences

UCLA Data Mining Lab **Los Angeles, CA**
GNNs in Predicting and explaining energy barriers of Metallic Glasses Mar 2023 - Present
Advisor: Yizhou Sun (Associate Professor)

- Led research on utilizing Graph Neural Networks (GNNs) to interpret impact of atomic structure of metallic glasses on its properties. **Co-1st author accepted at NeurIPS 2023 AI4Mat**
- Proposed and Implemented Symmetrized GNN (SymGNN) capable of handling invariance under any Orthogonal Transformation by aggregating over a learned distributions of O(3)
- Identified the primary challenge of task as the invariance of energy barriers under rotation and reflection, and more generally under all three dimensional orthogonal transformations
- Conducted Experiments and demonstrated SymGNN performs better compared to current baselines including GCN, EGNN, EGAT in terms of prediction accuracy, and outperform traditional material science method both in terms of accuracy and computation speed
- Extended GNNExplainer to regression setting and applied it to generate insightful explanations
- Investigating more statistically significant explanations results and how to trace back machine learning explanations back to physical intuition.

UCLA Computational Machine Learning Lab **Los Angeles, CA**
Data Attribution in Diffusion Models Mar 2023 - Present
Advisor: Cho-Jui Hsieh (*Associate Professor)*

- Led research on data attribution for diffusion models. **Co-1st author under-review at CVPR 2024**
- Designed and implemented Diffusion-TracIn by extending influence estimation based on loss gradient approximation, demonstrated its efficacy through auxiliary task such as outlier detection

- Identified bias in influence estimation induced by timesteps, and proposed Diffusion-ReTrac to address the issue through re-normalization techniques
- Designed experiments and evaluation metrics to highlight Diffusion-ReTrac successfully mitigates bias and outperforms in tasks including image source tracing and targeted influence estimation
- Studied and mastered generative models such as Autoregressive, Flow-Based, Score-Matching, Variational AutoEncoder (VAE), Generative Adversarial Network (GAN), and Diffusion Models
- Researched on existing instance-based interpretations in supervised and unsupervised settings

UCLA Computational and Applied Math REU **Los Angeles, CA**
Statistical Methods in Examining Efficiency of Gang Reduction Program Jun 2023 - Aug 2023
Advisor: Andrea Bertozzi (*Professor), Jona Lelmi (Hedrick Assistant Adjunct Professor)*

- Researched on efficacy of Los Angeles Gang Reduction Program using statistical methods. **Co-1ˢ author publication in IEEE 2023 Big Data DS4EIW**
- Applied Difference-In-Differences to rigorously analyze causal relationship between participation in program and subsequent reduction in risk of youths engaging in gang-related activities
- Proposed and implemented novel methods to model groups of youths as networks, integrating Linear-In-Means model to examine the efficacy of peer group networks cultivated through collective activities in reducing susceptibility of youths in gangs
- Proposed novel application of dynamic mode decomposition with control (DMDc) in studying the peer network effect by comparing the result with traditional dynamic mode decomposition (DMD)

Internship Experience & Projects

Vmware, Inc **Remote**
Software Development Intern (MTS intern) Jul 2022 - Sep 2022

- Constructed a feature merged into an integrated platform of VMware
- Utilized frameworks Spring Boot and Vue.js to implement backend and frontend interaction
- Initiated and designed innovative, beyond-assignment advanced functionalities, including but not limited to sorting algorithms, fuzzy query mechanisms, and caching systems, to enhance overall system performance and user experience.
- Won as 1-st place intern with Best Presentation award; thrived in fast paced high-tech company

Extracurricular Activities

UCLA Olga Radko Endowed Math Circle **Los Angeles, CA**
Assistant Instructor Sep 2022 - Present

- Organized and administered weekly lesson plans focusing on advanced materials such as Python programming, introduction to abstract algebra, graph theory, probability theory, set theory, cryptography, and combinatorics for middle school students
- Integrated hands-on activities and monitoring assessment to promote students' academic achievement, successfully facilitated engaging educational experience and individual progress

UCLA Course Reader **Los Angeles, CA**
Honors Linear Algebra Sep 2022 - Dec 2022

- In charge of grading assignments and writing homework solutions for 20+ students
- Monitored and analyzed trends shown in students' assignments to assist in teaching progress

Skill Sets

Python, Pytorch, Deep Graph Library (DGL), C/C++, Git, Javascript, Java, Haskell, Lisp, Linux