Hao Liu

McKelvey Hall 2010, One Brookings Drive, Saint Louis, MO 63130

📞 314-203-9875   ✉ haoliu4221@gmail.com   in linkedin.com/in/hao-liu-929016161/   🌐 haoliu-cola.github.io

Research Interest

I'm a **Ph.D. candidate** at Washington University in St. Louis advised by Prof. Yixin Chen. My research primarily focuses on **Graph Neural Networks (GNNs)**, **Large Language Models (LLMs)**, **Multimodal Learning**, **Few-shot Learning**, and **Contrastive Learning**. I am particularly interested in designing and developing graph foundation models that integrate LLMs to address tasks requiring multimodal understanding, prompt-based learning, and free-form output generation. My work involve the large-scale pretraining of graph foundation models using self-supervised learning paradigms and their fine-tuning for downstream tasks in both graph and language domains. Additionally, I am investigating tabular representation learning, with a particular focus on its applications to solving real-world clinical problems.

Education

Washington University in St. Louis   Saint Louis, MO
Ph.D. Candidate in Computer Science & Engineering   *Sep. 2019 – (Expected) May 2025*

Beijing Normal University   Beijing, China
B.S. in Mathematics   *Sep. 2015 – June 2019*

Selected Research Projects

GOFA: A Generative One-For-All Model for Joint Graph Language Modeling Feb. 2024 – Present
- Designed a graph foundation model that integrates GNN layers with LLM layers, leveraging the generative strengths of LLMs for free-form output and the structural learning capabilities of GNNs.
- Pioneered a novel pretraining paradigm focused on graph-level next-token prediction, facilitating large-scale self-supervised learning on various text-attributed graphs.
- Achieved state-of-the-art zero-shot performance on six node/link/graph-level datasets after instruction fine-tuning on a small number of data, demonstrating its potential as a foundation model.

One for All: Training One Graph Model for All Classification Tasks   Mar. 2023 – Jan. 2024
- Developed the first graph foundation model capable of handling multiple classification tasks across various datasets and domains in supervised, few-shot, and zero-shot scenarios.
- Employed Large Language Models to standardize and unify graph data representation, transforming features of diverse graphs into consistent natural language formats, enabling multi-dataset training in the graph domain.
- Innovated a graph prompting paradigm, advancing the model's in-context learning capabilities.

Unsupervised Meta-Learning for Transductive Few-shot Node Tasks   Jan. 2023 – Oct. 2023
- Designed the first unsupervised meta-learning method for transductive few-shot node classification.
- Innovated an unsupervised meta-task construction method by leveraging the strengths of contrastive learning, enabling comprehensive use of graph nodes in the learning process.
- Achieved state-of-the-art performance on seven datasets, with at least 11.18% and up to 20.56% absolute accuracy improvement over existing meta-learning baselines.

Publications

Conference:

- GOFA: A Generative One-For-All Model for Joint Graph Language Modeling
  L. Kong*, J. Feng*, **Hao Liu***, C. Huang, J. Huang, Y. Chen, M. Zhang
  The Thirteenth International Conference on Learning Representations (**ICLR 2025**)

- One for All: Towards Training One Graph Model for All Classification Tasks
  Hao Liu*, J. Feng*, L. Kong*, N. Liang, D. Tao, Y. Chen, M. Zhang
  The Twelfth International Conference on Learning Representations (**ICLR 2024 Spotlight**)

- Graph Contrastive Learning Meets Graph Meta Learning: A Unified Method for Few-shot Node Tasks
  Hao Liu, J. Feng, L. Kong, D. Tao, Y. Chen, M. Zhang
  The Web Conference 2024 (**WWW 2024**)

- TabContrast: A Local-Global Level Method for Tabular Contrastive Learning
  Hao Liu, Y. Chen, B. Fritz, C. King
  NeurIPS 2023 Second Table Representation Learning Workshop (**NeurIPS 2023 TRL**)

- MAG-GNN: Reinforcement Learning Boosted Graph Neural Network
  L. Kong, J. Feng, **Hao Liu**, D. Tao, Y. Chen, M. Zhang
  Thirty-seventh Conference on Neural Information Processing Systems (**NeurIPS 2023**)

- Extending the Design Space of Graph Neural Networks by Rethinking Folklore Weisfeiler-Lehman
  J. Feng, L. Kong, **Hao Liu**, D. Tao, F. Li, M. Zhang, Y. Chen
  Thirty-seventh Conference on Neural Information Processing Systems (**NeurIPS 2023**)

Preprint:

- TAGLAS: An Atlas of Text-attributed Graph Datasets in the Era of Large Graph and Language Models
  J. Feng, **Hao Liu***, L. Kong*, Y. Chen, M. Zhang

- Time Associated Meta Learning for Clinical Prediction
  Hao Liu, M. Zhang, Z. Dong, L. Kong, Y. Chen, B. Fritz, D. Tao, C. King

(* indicates equal contribution)

Technical Skills

Programming Languages: Python, Matlab, C
Machine Learning Frameworks: PyTorch, Lightning, PyG (Graph Neural Networks), Scikit-learn
Data Science Tools: Pandas, NumPy, Matplotlib, Seaborn
Developer Tools: PyCharm, Jupyter Notebooks, Git, Docker, Google Cloud Platform, VS Code

Work Experience

Applied Research Intern Capital One
Designed a graph language model that retrieves task-relevant information during message passing.June 2024 - Aug. 2024

Teaching Services

Washington University in St. Louis
CSE 543T: Nonlinear Optimization *Lecturer/Grader— Spring 2023/Spring 2024*