Yi (Owen) Yang

owen-yang-18.github.io | owen.yang@duke.edu | 470.943.9490

Education

Duke University, Trinity College of Arts and Science **Durham, NC**
Doctor of Philosophy in Computer Science Sep. 2023 - Present

- *Cumulative GPA:* 4.00/4.00
- *Teaching Assistantship:* Introduction to Databases, Fall 2024
- *Relevant Coursework:* Foundations of Blockchain, Theory and Algorithms of Machine Learning, Database Systems, Circuit Complexity, Graph Matrix Analysis

Emory University, College of Arts and Science **Atlanta, GA**
Bachelor of Science in Computer Science and Mathematics Aug. 2019 - May. 2023

- *Cumulative GPA:* 3.92/4.00; Dean's List: Fall 2021 & Spring 2022 & Fall 2022
- *Teaching Assistantship:* Data Structures and Algorithms, Fall 2021 & Fall 2022
- *Relevant Coursework:* Artificial Intelligence, Partial Differential Equations, Data Mining, Computer Architecture, Convex Optimization, Mathematical Statistics, Advanced Algorithms

Publication

Yang, Yi, Yanqiao Zhu, Hejie Cui, Xuan Kan, Lifang He, Ying Guo, and Carl Yang (2022). "Data-Efficient Brain Connectome Analysis via Multi-Task Meta-Learning". In: *the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining.*

Yang, Yi, Hejie Cui, and Carl Yang (2023). "PTGB: Pre-Train Graph Neural Networks for Brain Network Analysis". In: *the Conference on Health, Inference, and Learning.* (**Oral**).

Han, Kaiqiao, **Yang, Yi**, Zijie Huang, Xuan Kan, Yang Yang, Ying Guo, Lifang He, Liang Zhan, Yizhou Sun, Wei Wang, and Carl Yang (2024). "BrainODE: Dynamic Brain Signal Analysis via Graph-Aided Neural Ordinary Differential Equations". In: *Proceedings of the IEEE International Conference on Biomedical and Health Informatics.*

Li, Alexis, **Yang, Yi**, Hejie Cui, and Carl Yang (2024). "Brainsteam: A practical pipeline for connectome-based fmri analysis towards subject classification". In: *Pacific Symposium on Biocomputing.* (**Oral**).

Yang, Yi, Han Xie, Hejie Cui, and Carl Yang (2024). "Fedbrain: Federated training of graph neural networks for connectome-based brain imaging analysis". In: *Pacific Symposium on Biocomputing.*

Research Experiences

Duke University, Department of Computer Science **Durham, NC**
Research Assistant, Advisor: Jian Pei, Ph.D. Sep. 2023 - Present

- **Paradox Learning and Data Storytelling** *(On-going)*
  - Discovering that multiple instances of association reversal (*e.g.,* Simpson's paradox) in a multi-attribute dataset can describe the same data subspace – a phenomenon termed redundancy – and conjecturing that redundancy among association reversals establishes an equivalence relation.
  - Designing computational methods for efficient discovery of paradox redundancy by employing

pruning strategies to eliminate exhaustive and repetitive data subspace searches, and implementing parallelism and simplified data structures to improve execution and memory efficiency.
  - Constructing a synthetic data generator to uncover relevant conditions and properties of data distributions that precipitate formation of redundant instances of statistical paradoxes.

- **General Pattern Co-clustering of Sparse Tabular Data** *(On-going)*
  - Developing an unsupervised framework for co-cluster discovery in sparse tabular data by learning latent representations that model similarity structures between row and column objects.
  - Performing empirical studies on diverse real and synthetic datasets, providing interpretations to discovered co-clusters to support downstream applications such as recommendation systems.

Emory University, Department of Computer Science **Atlanta, GA**
Undergraduate Research Assistant, Advisor: Carl Yang, Ph.D. Aug. 2021 - Apr. 2023

- **Meta-learning for Brain Network Analysis** *(Published in KDD'22, PDF)*
  - Developed a meta-learning framework for graph neural networks, improving knowledge generalization across brain network systems and adaptation to local datasets with limited samples.
  - Introduced adaptive and task-specific learning trajectories to reconcile heterogeneous cross-tasks model performance due to varying task difficulties, resulting in faster model convergence and a 21% relative improvement in downstream performance over the strongest baseline.

- **GNN Pre-training for Brain Network Analysis** *(Published in CHIL'23, PDF)*
  - Proposed a self-supervised GNN pre-training for brain network data by optimizing a novel contrastive objective that leverages functional and spatial dependencies among cerebral regions, achieving a 13% relative increase in disease prediction accuracy over supervised baselines.
  - Designed a data-driven pre-processing pipeline to address the format misalignment issues in cross-dataset learning with a focus on the preservation of original brain parcellation structures.

- **Federated Learning for Brain Network Analysis** *(Published in PSB'24, PDF)*
  - Implemented a federated learning framework for brain network representation learning, enabling collaborative and distributed learning among multiple institutions with privately enclosed data.
  - Developed a hierarchical clustering of FL participants using clinical prior knowledge and neural circuitry metadata to enhance client personalization, leading to a 14% relative improvement in biomarker and disease prediction accuracy over state-of-the-art FL architectures.

Professional Experiences

C8 Technologies **Shanghai, China**
Portfolio Analyst Intern Jan. 2021 - Aug. 2021

- Prepared comprehensive reports on the fundamentals and mathematical intuitions of multiple industry-standard asset management strategies, utilized in company publications and marketing collateral.
- Redesigned the front-end interface of the company's online portfolio allocation toolbox and pioneered the implemention of a scalable backend for regression analysis and prediction on portfolio returns.

Services & Awards

Conf. & Jour. Reviews	TNNLS (2023, 2024), TKDD (2024), ICDM (2023, 2024), PSB (2024)
Awards & Honors	SIGKDD Student Travel Award (2022)