RENMING (REMY) LIU

4902 Forbes Ave Rm 7401, Pittsburgh, PA 15213

E-mail: renmingl@andrew.cmu.edu ⋄ Website: remylau.github.io ⋄ X: @RemyLau3

RESEARCH INTERESTS

- Network Biology
- Graph Representation Learning
- AI for Health & Science
- Single-Cell Analysis

EDUCATION

Michigan State University   09/2019 - 05/2024
Ph.D., Computational Mathematics, Science and Engineering

Michigan State University   09/2014 - 06/2018
B.S., Chemical Engineering
Concentration in Biochemical Engineering
Minor in Computer Science

PROFESSIONAL EXPERIENCE

Post-doctoral Fellow   06/2024 - present
Ray and Stephanie Lane Computational Biology Department. Carnegie Mellon University.
Topic: Machine learning methods and software for single-cell epigenomics.

Ph.D. Student and Research Assistant   09/2019 - 05/2024
Department of Computational Mathematics, Science and Engineering. Michigan State University.
Topic: Biomedical and Molecular Interaction Networks, Graph Representation Learning, Spectral Graph Theory.

Research Assistant   06/2023 - 08/2023
Department of Biomedical Informatics. Colorado University Anschutz Medical Campus.
Topic: Context-specific biological network embeddings.

Research Assistant   05/2017 - 08/2019
Department of Chemical Engineering and Materials Science. Michigan State University.
Topic: Thermodynamic modeling of associating systems including alcohols and acetic acids.

Professional Aide   09/2018 - 08/2019
Department of Computational Mathematics, Science and Engineering. Michigan State University.
Topic: Human gene function prediction using ML with biological networks.

HONORS & AWARDS

Silver Medal in Kaggle   2022
NeurIPS Multimodal Single-Cell Integration Across Time, Individuals, and Batches.

Outstanding Graduate Student Award   2022
Awarded to the most outstanding doctoral student in each program.

Raymond P. and Marie M. Ginther Fellowship   2019
Awarded to outstanding incoming Ph.D. students to the CMSE program at MSU.

Maurice G. and Sara V. Larian Scholarship   2017

OPEN-SOURCE PROJECTS

OBNB (sole developer) 2023
A Python toolkit for benchmarking biomedical networks using gene annotation resources.

DANCE (core developer) 2022
An extensive deep learning toolkit for single-cell (multi-)omics data.

PyGenePlexus 2022
A network-based gene classification service using ML with gene interaction network features.

PecanPy (sole developer) 2020
A memory efficient and parallel Python implementation of node2vec with an improved version *node2vec+* for weighted graph.

PUBLICATIONS

indicates (co-)first author

Journal and Conference Papers

- Graph Positional and Structural Encoder
  S. Cantürk*, **R. Liu***, O. Lapointe-Gagné, V. Létourneau, G. Wolf, D. Beaini, L. Rampášek
  International Conference on Machine Learning (ICML), 2024

- DANCE: A Deep Learning Library and Benchmark Platform for Single-cell Analysis
  J. Ding*, H. Wen*, W. Tang*, **R. Liu***, Z. Li, J. Venegas, R. Su, D. Molho, W. Jin, W. Zou, Y. Wang, R. Yang, Y. Xie, J. Tang
  Genome Biology, 2024

- Deep Learning in Single-Cell Analysis
  D. Molho*, J. Ding, Z. Li, H. Wen, W. Tang, Y. Wang, J. Venegas, W. Jin, **R. Liu**, R. Su, P. Danaher, R. Yang, Y. Lei, Y. Xie, J. Tang
  ACM Transactions on Intelligent Systems and Technology, 2024

- Joint representation of molecular netowrks from multiple species improves gene classification
  C.A. Mancuso*, K.A. Johnson, **R. Liu**, A. Krishnan
  PLOS Computational Biology, 2024

- SpatialCTD: a large-scale TME spatial transcriptomic dataset to evaluate cell type deconvolution for immuno-oncology
  J. Ding*, J. Venegas, Q. Lu, Y. Wang, W. Jin, H. Wen, **R. Liu**, W. Tang, Z. Li, W. Zuo, Y. Chang, Y. Lei, P. Danaher, Y. Xie, J. Tang
  Journal of Computational Biology, 2024

- Open Biomedical Network Benchmark: A Python Toolkit for Generating Benchmakring Datasets with Biomedical Networks
  R. Liu*, A. Krishnan
  Machine Learning in Computational Biology (MLCB). PMLR, 2023

- Single-Cell Multimodal Prediction via Transformers
  W. Tang*, H. Wen*, **R. Liu***, J. Ding, W. Jin, Y. Xie, H. Liu, J. Tang
  ACM International Conference on Information & Knowledge Management (CIKM), 2023

- Accurately modeling biased random walks on weighted networks using *node2vec+*
  R. Liu*, M.J. Hirn, A. Krishnan
  Bioinformatics, 2023

- PyGenePlexus: A Python package for gene discovery using network-based machine learning
  C.A. Mancuso*, **R. Liu**, A. Krishnan
  Bioinformatics, 2023

- Taxonomy of Benchmarks in Graph Representation Learning
  R. Liu*, S. Cantürk*, F. Wnkel, S. McGuire, X. Wang, A. Little, L. O'Bray, M. Perlmutter, B. Rieck, M.J. Hirn, G. Wolf, L. Rampášek
  Learning on Graph (LoG) Conference, 2022

- GenePlexus: A web-server for gene discovery using network-based machine learning
  C.A. Mancuso*, P.S. Bills, D. Krum, J. Newsted, **R. Liu**, A. Krishnan
  Nucleic Acids Research, 2022

- Applications of an Association Activity Coefficient Model, NRTL-PA, to Alcohol-Containing Mixtures
  A.M. Bala*, **R. Liu**, L. Peereboom, C.T. Lira
  Industrial & Engineering Chemistry Research, 2022

- PecanPy: a fast, efficient and parallelized Python implementation of node2vec
  R. Liu*, A. Krishnan
  Bioinformatics, 2021

- Supervised Learning is an accurate method for network-based gene classification
  R. Liu*, C.A. Mancuso*, A. Yannakoppoulos, K.A. Johnson, A. Krishnan
  Bioinformatics, 2020

Preprints and Submissions

- A General Single-Cell Analysis Framework via Conditional Diffusion Generative Models
  W. Tang*, **R. Liu***, H. Wen, X. Dai, J. Ding, H. Li, W. Fan, Y. Xie, J. Tang
  bioRxiv

- CONE: COntext Specific Network Embedding via Contextualized Graph Attention
  R. Liu*, H. Yuan*, K.A. Johnson, A. Krishnan
  Submitted to MLCB'24

- Memory-Augmented Pretraining for Single-Cell Multi-Omic Analysis
  X. Dai*, H. Wen, W. Tang, **R. Liu**, Y. Wu, J. Ding, W. Jin, L. Yu, Y. Xie, J. Tang
  Submitted to IDCM'24

- Single Cells Are Spatial Tokens: Transformers for Spatial Transcriptomic Data Denoising
  H. Wen*, W. Tang, W. Jin, J. Ding, **R. Liu**, F. Shi, Y. Xie, J. Tang
  arXiv

PRESENTATIONS

Invited Talks

- Professor Trey Ideker's lab at University of California San Diego   02/2024

- Professor Jian Ma's lab at Carnegie Mellon University   02/2024

- Learning on Graphs (LoG) local meetup at Mid-North US   01/2024

Oral Presentations

- Systematic study of human diseases using graph signal processing   12/2022
  ISCB-Rocky'22

- Taxonomy of Benchmarks in Graph Representation Learning 12/2022
  LoG'22

- Understanding dataset characteristics via graph diffusion 09/2022
  SIAM-MDS'22

- Accurately identify disease genes and contexts using context-specific network embeddings 07/2022
  ISMB'22

- Accurate network-based gene classification with ultra-fast context-specific node embeddings 05/2021
  GLBIO'21

Posters

- Open Biomedical Network Benchmark: A Python Toolkit for Generating Benchmakring Datasets with Biomedical Networks 11/2023
  MLCB'23

- A Python Toolkit for Generating Benchmarking Datasets for Machine Learning with Biological Networks 03/2023
  CSHL-NetBio'23

- PecanPy: a fast, efficient, and parallelized Python implementation for node2vec 07/2020
  ISMB'20

- Supervised-learning is an accurate method for network-based gene classification 08/2019
  CoCoH'19

- Machine learning outperforms label propagation for network-based gene classification 03/2019
  CSHL-NetBio'19

SERVICES

Teaching

- Teaching Assistant for CMSE 201: Introduction to Computational Modeling 2020
  Lecture assistant, office hours, and grading.

- Grader for CHE 804: Foundations in Chemical Engineering 2019

- Grader for CHE 321: Thermodynamics for Chemical Engineering 2019

Mentoring

- Keenan Manpearl (she/her) 03/2023 - Present
  Ph.D. student at University of Colorado Anschutz Medical Campus
  Ongoing Project: Leveraging ontology embeddings to enable transfer learning for gene classification.

Conference & Workshop Program Chairs

- The Web Conference 2024 Graph Foundation Model Workshop 2024
  Program Chair

Journal and Conference Paper Reviews

- Machine Learning in Computational Biology (MLCB) 2024

- NeurIPS Datasets and Benchmarks Track 2024

- Transactions on Knowledge and Data Engineering (TKDE) 2023 - 2024

- Transactions on Knowledge Discovery from Data (TKDD) 2023 - 2024

- F1000 Research   2023
- Cell Systems   2022
- Nature Computational Science   2022
- Bioinformatics   2021
- eLife   2019