HONGKANG LI

+1(518) 961-2812 ◇ Troy, NY
lih35@rpi.edu ◇ Google scholar

EDUCATION

Rensselaer Polytechnic Institute, Troy, NY, United States August 2019 - Present
P.h.D. candidate, Electrical and Computer System Engineering
GPA: 4.0/4.0
Advisor: Prof. Meng Wang

University of Science and Technology of China, Hefei, China 2015 - 2019
Bachelor of Engineering, GPA: 3.73/4.3

RESEARCH INTEREST

Machine Learning Theory: deep learning theory, generalization analysis of (graph) neural networks, theoretical mechanism of large language models.
Trustworthy Machine Learning: model explanation, fairness
Data Science: data-efficient learning

SELECTED RESEARCH PROJECTS

Training Transformers for Chain-of-Thought Inference

The contribution of this work includes (a) quantifying how the training of Transformers can enable the Chain-of-Thought ability in inference; (b) characterizing how the quality of contexts affects the Chain-of-Thought performance; (c) studying the comparison between Chain-of-Thought and In-Context Learning.

Generalization of In-Context Learning with Transformers

This work (a) proposes an analytical framework for training and in-domain and out-of-domain generalization of Transformers for in-context learning; (b) characterizes the mechanism of in-context learning using Transformers; (c) theoretically studies the validity of model pruning.

Generalization Analysis of Graph Transformers on General Graphs

This project focuses on the following aspects. (a) Optimization and generalization analysis of Graph Transformers. (b) Comparisons between generalization performance of Graph Transformers. and GCNs. (c) The effect of using positional encoding to promote the core neighborhood.

Sample Complexity Analysis of Learning with Vision Transformers

This work (a) proposes a new analytical framework to tackle the non-convex optimization and generalization for shallow ViTs; (b) theoretically depicts the evolution of the attention map during the training; (c) provides a theoretical explanation for the improved generalization using token sparsification.

SELECTED PUBLICATIONS

* denotes equal contribution

H. Li*, M. Wang, S. Lu, X. Cui, and P. Chen. "How Do Nonlinear Transformers Acquire Generalization-Guaranteed CoT Ability?" *ICML 2024 Workshop on Theoretical Foundations of Foundation Models; ICML 2024 Workshop on High-dimensional Learning Dynamics: The Emergence of Structure and Reasoning*

Y. Zhang*, **H. Li*,** Y. Yao*, A. Chen, S. Zhang, P. Chen, M. Wang, S. Liu, "Visual Prompting Reimagined: The Power of Activation Prompts." *In submission.*

H. Li, M. Wang, S. Lu, X. Cui, and P. Chen. "How Do Nonlinear Transformers Learn and Generalize in In-Context Learning?" *Accepted by International Conference on Machine Learning (ICML)(**acceptance rate = 27.5%**), 2024.*

H. Li, M. Wang, T. Ma, S. Liu, Z. Zhang, P. Chen, "What Improves the Generalization of Graph Transformer? A Theoretical Dive into Self-attention and Positional Encoding. " *Accepted by International Conference on Machine Learning (ICML)(**acceptance rate = 27.5%**), 2024.*

S. Zhang, **H. Li,** Meng Wang, Miao Liu, Pin-Yu Chen, Songtao Lu, Sijia Liu, Keerthiram Murugesan, Subhajit Chaudhury. "On the Convergence and Sample Complexity Analysis of Deep Q-Networks with -Greedy Exploration. " *Accepted by Annual Conference on Neural Information Processing Systems (Neurips) (**acceptance rate = 26.1%**), 2023.*

H. Wan, **H. Li,** Songtao Lu, Xiaodong Cui, Marina Danilevsky. "How Can Personalized Context Help? Exploring Joint Retrieval of Passage and Personalized Context. "*Accepted by IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2023.*

H. Li, S. Zhang, M. Wang, Y. Zhang, P. Chen, S. Liu, "Theoretical characterization of neural network generalization group imbalance." *Accepted by IEEE Journal of Selected Topics in Signal Processing.*

H. Li, M. Wang, S. Liu, P. Chen, "A theoretical understanding of shallow vision transformers: Learning, generalization, and sample complexity." *Accepeted by International Conference on Learning Representations (ICLR)(**acceptance rate = 32.0%**), 2023.*

H. Li, M. Wang, S. Liu, P. Chen, and J. Xiong. "Generalization guarantee of training graph convolutional networks with graph topology sampling." *Accepted by International Conference on Machine Learning (ICML)(**acceptance rate = 21.9%**), 2022.*

H. Li, S. Zhang, and M. Wang. "Learning and generalization of one-hidden-layer neural networks, going beyond standard gaussian data." *Accepted by 2022 56th Annual Conference on Information Sciences and Systems (CISS). IEEE*

SERVICES

- Reviewer of
  - ICML 2022-2024;
  - Neurips 2022-2024;
  - ICLR 2024-2025;
  - CVPR 2024;
  - AAAI 2025.
  - AISTATS 2024.
  - UAI 2023;
  - ICASSP 2023-2024;
  - TMLR 2023-2024.
- Mentor of research of three ECSE undergraduate students of RPI.

HONORS AND AWARDS

- **Founders Award of Excellence**, Rensselaer Polytechnic Institute — October 2023
- **Belsky Award**, School of Engineering, Rensselaer Polytechnic Institute — January 2023
- **Zenghua Scholarship**, Top 5%, University of Science and Technology of China — September 2018
- **Outstanding Student Scholarship**, University of Science and Technology of China — 2016 - 2017