Yun'ai Li

Contact: yl13580@nyu.edu | Homepage: https://li-yunai.github.io
Jersey City, NJ, 07302, US

Education and Research Affiliations

- **Shanghai Jiao Tong University (SJTU)** *Sep.2021-Jul.2025 (Expected)*
  Bachelor of Science In Mathematics and Applied Mathematics (Honor Track) Shanghai, China
  - Zhiyuan Honors Program, top 15 % in the School of Mathematical Sciences of SJTU
  - GPA:3.41/4.0

- **New York University, Center for Data Science (NYU CDS)** *Jul.2024-Jan.2025*
  Visiting Student for Undergraduate Research New York, NY
  - Advisor: Prof.Qi Lei

- **Nanyang Technological University, School of Physical and Mathematical Sciences** *Jan.2024-Jun.2024*
  Visiting Student for Undergraduate Research Singapore
  - Advisor: Prof.Juan-Pablo Ortega

Academic Interests

My primary research interests lie in Machine Learning theory and Optimization , with a particular focus on developing computationally efficient algorithms with theoretical guarantees.
I'm enthusiastic about exploring the related areas and would be excited to discuss about further opportunities (for 25fall phd) if you see the potential!

Research Experience * click on the titles to see the research details (reports, slides, code, etc)

- **Theoretical Framework and Provable Guarantee for Weak-to-strong Generalization** *New York*
  Keywords: *Weak supervision; data distillation; learning theory; Representation Learning* Ongoing
  Advisor: Prof.Qi Lei, New York University
  - Developing a theoretical framework for the Open AI's cutting edge paper on weak2strong generalization. Specifically focusing on giving the weak-supervised strong student model's quantifiable gain on various kinds of downstream tasks.
  - Providing theoretical guarantees for future weak-supervised models (e.g., RLHF, learning under restricted data access), offering new insights into data distillation and enhancing understanding of weak supervision.

- **Investigating Optimizer-Induced Implicit Bias in Transformers for NLP Tasks** *Remote*
  Keywords: *Self-attention Mechanism; Next Token Prediction; Adam/AdamW; optimization* Ongoing
  Advisor: Prof.Yingbin Liang, Ohio State University **Prof.Jing Yang**, Pennsylvania State University
  - Extending current work on implicit biases in transformers by providing theoretical analyses of training dynamics in specific NLP tasks, such as next token prediction
  - Investigating optimizer-induced biases of Adam/AdamW, moving beyond empirical observations to provide theoretical insights into their advantages and efficiency in transformer models.
  - Relaxing existing assumptions in the current work to broaden the theoretical understanding of transformer training, with a focus on practical implications in NLP tasks.

- **Innovative Reservoir Computing Approaches for Reinforcement Learning** *Singapore*
  Keywords: *Reservoir computing; Echo state network; Actor-critic algorithm, learning theory* Jan-May, 2024
  Advisor: Prof.Juan-Pablo Ortega, Nanyang Technological University
  - Integrated Echo State Networks (ESNs) into reinforcement learning frameworks (policy-based, value-based, and actor-critic) to develop novel algorithms, leveraging ESNs' universal approximation capabilities to improve performance and generalization on complex tasks.
  - Enhanced ESN-based algorithms with advanced reinforcement learning techniques (e.g., PPO, SAC, DQN), outperforming traditional architectures (LSTM, MLP) in terms of optimization stability and gradient-free updates.
  - Explored the general least-squares TD (LSTD) algorithm through the lens of ESNs, examining how ESNs' dynamics can optimize TD learning, enhance sample efficiency, and address stability challenges in value-based reinforcement learning.
  - Conducted rigorous asymptotic and finite-time analyses of the Actor-Critic with ESN algorithms, demonstrating superior convergence rates and sample efficiency, highlighting ESNs' inherent regularization benefits in complex environments.

SELECTED COURSEWORK AND PROJECTS* CLICK ON THE TITLES TO SEE PROJECT DETAILS

- **Selected Course Projects:**
  - Generative Models for Aircrafts' Icing Image Prediction
    Neural Encoding in Balanced Networks: Data-Driven Exploration
    PINNs for Helmholtz Equations' Forward/Inverse Problems in Multiple Propagation Mediums
    Multinomial Logit Modeling and Numerical Optimization for Air Ticketing Behavior Analysis
- **Selected Course Essays:**
  - *Exterior Algebra, Hypergraph and Some Relative Inequalities about the Bollobas Two Families Theorem // A Review on C. E. Shannon's Original Paper and Discussions about Graph Entropy // Classification of 18-Element Groups // The Solution and Properties of Three-Dimensional Linear Systems with Constant Coefficients*
- **Selected Advanced Courses:**
  - Stochastic Process(Analysis), Foundations of Data Science, Graph theory, Mathematical Programmings
  - PDE, Real Analysis, Dynamic System (Ergodic Theory), Differential Geometry, Abstract Algebra (with Galois Theory), Functional Analysis (expected)

HONORS AND AWARDS

- **Scholarships and Grants:**
  - Zhiyuan Outstanding Research Visiting Fund (for visiting CAS) *(2023)*
  - SJTU Outstanding Undergraduate Student Scholarship (top 15%) *(2022)*
  - Zhiyuan Honorary Scholarship (top 5%) *(2022/2021)*
- **Miscellaneous:**
  - Mathematical Contest in Modeling: Honorable Mention(MCM/ICM) *(2023)*

SKILLS

- **Languages:** Chinese (Native), English (Fluent, IELTS 7.5, TOEFL 103)
- **Coding skills:** Python (PyTorch, Tensorflow), MATLAB, LATEX, Markdown, HTML