Lize Shao

Houston ls94@rice.edu 346 227 0075

Education

Rice University *Aug 2021 – May 2025*
Dual Degree: BA in Mathematics, BS in Computer Science *MAJOR GPA: 3.80*

- **4-year Roy Trustee Distinguish Scholarship (TOP1% Based on Holistic Evaluation)**
- **Coursework:** Data Structure and Algorithm, Concurrent&Parallel PROG, Computer Network&Architecture, Mobile&Embedded Sys, Machine Learning, Compiler Construction, Calculus&Linear Algebra, Combinatorics, Number Theory, Real Analysis, Topology, Differential Equation, Micro/Macro/Behavioral Economics
- **Test Scores:**
  - **GRE:** 336; Verbal Reasoning (166), Quantitative Reasoning (170), Analytical Writing (4.5)

Experience

Student Researcher *Houston, TX*
Supervised by Xia Hu, DATA Lab, Rice University *Aug 2023 – Sep 2024*

- Led the comprehensive audit and enhancement of the MQuAKE dataset, addressing critical flaws and redefining standards for multi-hop knowledge editing.
- Designed and implemented solutions including the dynamic masking and GWalk algorithms to advance state-of-the-art LLM knowledge editing.
- Co-authored one paper, currently under review at ICLR 2025, which introduces scalable, high-precision frameworks for multi-hop editing tasks.

Lead Researcher *Houston, TX*
Supervised by Xia Hu, DATA Lab, Rice University *Mar 2024 – Present*

- Developed a hybrid AI-driven system combining LLMs with neural-symbolic reasoning to automate fault detection and performance tuning in large-scale distributed environments.
- Designed scalable algorithms for dynamic resource allocation and fault-tolerant execution.
- Introduced a graph-theoretic technique to model and analyze complex interdependencies in distributed architectures, enhancing system diagnostics and optimization workflows.
- Research outcomes with findings under preparation for submission to ICSE 2025.

Research Intern *Remote*
Supervised by Tegawendé F.BISSYANDE, TruX Lab, University of Luxembourg *Sep 2023 – Present*

- Designed a hyperbolic model, analyze its ability to capture hierarchical relationships between code and natural language.
- Reframed traditional code retrieval tasks into a representation learning framework, achieving a 3.5%-4% improvement over state-of-the-art code search models.
- Published one paper on leveraging hyperbolic geometry to enhance the performance of code retrieval systems.

Lead Researcher *Remote*
Supervised by Tegawendé F.BISSYANDE, TruX Lab, University of Luxembourg *June 2024 – Present*

- Led the development of AnyCoding, an AI-powered autonomous programming system addressing limitations of existing single-file repair tools by employing dynamic call graphs for holistic code repair.
- Designed advanced routing algorithms and intelligent agents that resolve interdependent issue resolution across variables, methods, classes, and filenames.
- Achieved 32% resolved rate on the SWE-bench-lite dataset and 73% resolved rate on 15 full-scale code projects. Co-authored a paper, which is currently under review at ACL 2025.

Teaching Assistant *Houston, TX*
Supervised by Michael Burke, The Department of Computer Science, Rice University *Aug 2023 – Present*

- Organize, lead, and supervise laboratory sessions. Guide students on applications of data structures and algorithms.
- Developed lecture presentations and created exam and homework questions to support course instruction.

- Designed grading rubrics and evaluated assignments, exams, and projects.

Software Engineer Intern *Beijing, China*
Mentored by Jiameng Huang, R&D Department, Microsoft Corp. *June 2023 – Aug 2023*

- Engineered an end-to-end solution to optimize Microsoft Teams metadata processing, improving search relevance and engagement metrics for 150M+ users globally.
- Collaborated with cross-functional teams, including product managers and UX designers, to enhance telemetry logging for user interaction analytics and feature usage monitoring.
- Designed and implemented scalable data pipelines using Python, Azure Data Factory, and Azure Synapse Analytics, ensuring high reliability and real-time data processing capabilities.
- Developed and deployed an interactive Power BI dashboard to visualize key performance indicators, empowering stakeholders to make data-driven decisions and improve system performance.

Machine Learning Engineer Intern *Shanghai, China*
Mentored by Zhendong Wang, Intelligent Creation Lab ByteDance Corp. *May 2022 – Aug 2022*

- Engaged in deep learning techniques to develop an image captioning system that describes input images with relevant text.
- Utilized a CNN encoder, adopting the pre-trained InceptionV3 model. Implemented a GRU-based decoder for the caption generation process.
- Spearheaded algorithmic development using Python and transitioned to backend interface coding using Java in the project's latter stages.

Projects

BESSER-PEARL: A Low-Code Open Source Environment for Modeling Complex Systems

- Contributed to the development of a low-code platform enabling users to model, simulate, and analyze complex systems.
- Designed intuitive workflows and developed features to support both graphical and textual modeling paradigms.

AR Guided TBP Capturing: Specifications and Requirements Documentation

- Developed an AR-guided mobile application for Total Body Photography (TBP) to aid in early skin cancer detection while ensuring privacy compliance.
- Implemented AR-guided positioning, multi-angle photography, and automated blurring for patient confidentiality.
- Designed features for patient data management, doctor collaboration, and consent-based data sharing.
- Built a scalable backend with Flask and MongoDB, integrated AWS S3 for photo storage, and ensured HIPAA compliance.

OwlDb Database Web Service

- Designed and implemented a RESTful web service using Go to provide a robust, network-accessible NoSQL document database. Enabled CRUD operations for JSON documents, allowing users to create, modify, retrieve, and delete entries effortlessly.
- Integrated a real-time observation mechanism for monitoring changes in documents, delivering instant updates to clients.
- Implemented efficient indexing and search mechanisms, optimizing performance for high-volume data access and operations. Leveraged the HTTP protocol for seamless integration with two platforms.

Publications and Papers

HoCoS: Hyperbolic Representation Towards Code Search

Xunzhu Tang*, ***Lize Shao****, Yewei Song, Saad Ezzini, Haoye Tian, Jiechao Gao, Jacques Klein, Tegawendé F Bissyandé

MQuAKE-Remastered: Multi-Hop Knowledge Editing Can Only Be Advanced With Reliable Evaluations

Shaochen Zhong, Yifan Lu, ***Lize Shao***, Bhargav Bhushanam, Xiaocong Du, Louis Feng, Yixin Wan, Yucheng Shi, Daochen Zha, Yiwei Wang, Ninghao Liu, Kaixiong Zhou, Shuai Xu, Vipin Chaudhary, Xia Hu

AnyCoding: An Autonomous Artificial Intelligent Programmer

Lize Shao*, Xunzhu Tang*, Jiechao Gao, Haoye Tian, Bach D X Le, Jacques Klein, Tegawendé F. Bissyandé

FALCON: A Hybrid AI Framework for Fault Detection and Resource Optimization in Distributed Systems

Lize Shao, Alan, Cao*, Yifan Lu*, Mike Wei, Henry Zhong, Xia Hu

Zero-Shot Stance Detection Enhanced with Augmented Background Knowledge Based on LLMs

Lize Shao, Jacky Jiang*, Jerry Wei*, Vicente Ordóñez-Román

Fed-UGen: Uncertainty-Guided Federated Learning for Domain Generalization

Lize Shao*, Creed Gao*

Awards

- **Roy Trustee Distinguished Scholarship** — TOP 1% Based on Holistic Evaluation 2021-Present
- **Rice Datathon** — 2nd Place in the Chevron Challenge 2023
- **USA Computing Olympiad** — Global TOP 52 2021
- **USA Math Olympiad** — Qualified for USAMO (Global Top 100) 2020
- **Canadian Math Olympiad** — Global Gold Award 2020