YINING SHE

Pittsburgh, PA yiningsh@cs.cmu.edu sheyining.github.io

EDUCATION

Carnegie Mellon University *Pittsburgh, PA*
Ph.D. in Software Engineering, Advised by Eunsuk Kang *Aug.2022-May 2027(expected)*
Software and Societal Systems Department, School of Computer Science
Research focus: Software Engineering for AI; Robustness, Fairness & Reliability of AI System

ShanghaiTech University *Shanghai, China*
Bachelor of Engineering in Computer Science and Technology *Sep.2018-Jun.2022*
GPA: **3.83/4.0**
Selected Honors: *Outstanding Graduate, Merit Student, Outstanding Teaching Assistant*

PROFESSIONAL SKILLS

Programming Languages	Python, C/C++, C#, R, MATLAB
Tools and Frameworks	PyTorch, scikit-learns, Tensorflow, OpenCV, Alloy, UPPAAL, Git, RISC-V

DOCTORAL RESEARCH

Enhancing LLM Agent Safety Through Robust Toolkit Design *In-process*

- Proposing a novel framework to systematically design and implement safer toolkits for LLM agents to minimize unintended consequences such as data leakage or financial harm
- Developing automated checks and usage constraints within the toolkit to guide LLM agents toward safe operation, analogous to how well-engineered software guides human users
- Incorporating fine-grained safeguards into each tool, enabling strict control over high-impact actions (e.g., financial transactions or data-sharing requests)
- Introducing high-level, task-specific tools that encapsulate multi-step operations (like verifying user identities and processing payments) into a single, controlled workflow, reducing the risk of misuse by LLMs
- Planning to evaluate the robustness and practicality of the proposed design in real-world agents, measuring reductions in unsafe behaviors compared to baseline toolkits
- Complementary to existing approaches that focus on improving LLM models themselves, this research aims to provide a developer-centric toolkit framework that guarantee overall agent safety

FairSense: Long-Term Fairness Analysis of ML-enabled Systems *In ICSE 2025*

- Proposed a simulation-based framework FairSense to detect and analyze the long-term unfairness in ML-enabled systems
- Modeled feedback loop between an ML-enabled system and its deployment environment to evaluate fairness over time via Monte Carlo simulation.
- Performed sensitivity analysis on simulation traces to understand the impact of design options and environmental factors on the long-term fairness of the system
- Applied sampling heuristic to efficiently explore exponentially large configuration space without affecting sensitivity analysis results
- Evaluated FairSense on three real-world benchmarks: loan approval, opioid risk scoring, and predictive policing, and results show it can effectively detect long-term fairness issues and identify system configuration variables with the greatest impact for developers

INDUSTRY EXPERIENCE

Intel Lab *Portland, OR*
AI Software Research Intern, Parallel Computing Lab *May 2024-Aug.2024*

- Proposed an ML-based tensor kernel loop optimization method to predict the optimal loop configuration for a given machine
- Collected General Matrix Multiply (GEMM) kernel performance data across 1,040 loop tiling and ordering configurations on 15+ machines with diverse architectures and platforms

- Trained a multi-layer perceptron (MLP) model to predict kernel performance based on loop configuration and machine characteristics
- Integrated the ML model into a progressive ranked exploration pipeline that alternates between exploring new configurations and exploiting the current optimal solution
- Evaluated time efficiency of progressive ranked exploration compared to the baseline method, and results show it identifies the global optimal GEMM loop configuration for all 15+ machines by searching $\leq 3\%$ of the configuration space

PUBLICATIONS

- **Yining She**, Sumon Biswas, Christian Kästner, Eunsuk Kang. **"FairSense: Long-Term Fairness Analysis of ML-Enabled Systems"**, In the 47th IEEE/ACM International Conference on Software Engineering (ICSE), 2025.
- Sumon Biswas, **Yining She**, Eunsuk Kang. **"Towards Safe ML-based Systems in Presence of Feedback Loops"**, *In SE4SafeML, Proceedings of the ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE), 2023.*
- Mohammad Piran, **Yining She**, Renzhi Tang, Zhihao Jiang, Yash Vardhan Pant. **"Stable Interaction of Autonomous Vehicle Platoons with Human-Driven Vehicles"**, *In American Control Conference (ACC), 2022.*

TEACHING EXPERIENCE

TA for graduate course "Formal Methods"	*Carnegie Mellon University, Fall 2023*
Lead TA for undergraduate course "Algorithms and Data Structure"	*ShanghaiTech University , Fall 2021*
Lead TA for undergraduate course "Software Engineering"	*ShanghaiTech University , Spring 2021*
TA for undergraduate course "Algorithms and Data Structure"	*ShanghaiTech University , Fall 2020*