Xiyuan Yang

Phone: (+1) 217 766 5798 | **E-mail:** xiyuany4@illinois.edu | Google Scholar

EDUCATION

University of Illinois, Urbana-Champaign (US) 2025-2030
(First Year) PhD of Computer Science | Advisor: Jingrui He
Current Research Area: **Theory & Optimization of Foundation Models ; Trustworthy Machine Learning**

Wuhan University (China) 2021-2025
Bachlor of Computer Science | Advisor: Mang Ye | GPA 3.77 / 4.0

PUBLICATIONS (*bold name represents first author or equal contribution; the mentioned year is the time that the work has been finished)

[1] **Xiyuan Yang** et al., Preconditioning Neural Tangent Kernel for Adaptive Optimization, in submission of AISTATS, 2025.
[2] **Xiyuan Yang** et al., Differentially Private Federated Clustering with Random Rebalancing, in submission of ICLR, 2025.
[3] **Xiyuan Yang** et al., Defending against Indirect Prompt Injection by Instruction Detection, ACL, 2025.
[4] **Xiyuan Yang** et al., Defending LLMs Against Jailbreak Attacks Utilizing Cross-Modality Generalization Gap, in submission of Nature Communications, 2025.
[5] Chenglong Wang et al., Uncovering inequalities in new knowledge learning by large language models across different languages, in submission of PNAS, 2025.
[6] Yueqi Xie et al., Measuring human contribution in ai-assisted content generation, in submission of PNAS, 2024.
[7] **Xiyuan Yang** et al., FedAS: Bridging Inconsistency in Personalized Federated Learning, CVPR, 2024.
[8] Xiuwen Fang et al., Robust heterogeneous federated learning under data corruption, ICCV, 2023.
[9] **Xiyuan Yang** et al., Dynamic personalized federated learning with adaptive differential privacy, NeurIPS, 2023.

RESEARCH EXPERIENCE

Research Intern at University of Chicago Chicago, US
Advisor : Tian Li 2024-2025
Research Direction : Quantiative Privacy Leakage & Protection; Optimization on Long-tail Distributions **[2]**

Research Intern at Microsoft Research (Asia) Beijing, China
Advisor : Fangzhao Wu 2024-2025
Research Direction : Security Topics & Societal Impacts of Foundational Models (LLM, VLM, etc.) **[3, 4, 5, 6]**
Project Direction : Trained News Recommendation Models for MS News Group

Research Intern at Wuhan University Wuhan, China
Supervisor: Mang Ye 2022-2024
Research Direction : Distributed Optimization; Trustworthy Machine Leaning **[7, 8, 9]**

ACADEMIC SERVICE

Reviewer of Conferences (CVPR, ICCV, ICML, NeurIPS, ICLR, AAAI, etc.) and Journal (Inf Fusion, IEEE TKDE, TNNLS, etc.)
IEEE Student Member 2023-2025

SCHOLARSHIP

Overseas exchange and study scholarship of Wuhan Univeristy (Top 5%) 2024
LeiJun Computer Research Funding Scholarship (Top 0.5%) 2024

COMPETITION

MIND News Recommendation Competition (Rank 1/112 groups) 2024
- Incorporated pretrained LMs as the news encoder and user encoder, and achieved SOTA performance.

PROJECT EXPERIENCE

Chatbot Design Based on LLaMA-33B 2023
- Fine-tuned the pre-trained LLaMA model on open-source Chinese corpus to improve its ability in Chinese conversations.
- Used QLoRA technology to greatly reduce Video Memory usage, enabling fine-tuning on a single 3090.

CPU Design for RISC-V Instruction Set 2023
- Used the Verilog language to design and implement a five-stage pipeline CPU, including IF/ID/EX/MEM/WB stages.
- Implemented the decoding and execution of the RISC-V instruction set, including arithmetic, logic, load/store, branch, etc.

SKILLS

Technical: Proficient in Python, PyTorch (especially for deep learning based, LLM related code), Skilled in C/C++; Familiar with Java, common front-end technologies, Haskell functional programming language, and MySQL database