Bingqi Shang

Website | Google Scholar | GitHub | bshang@msu.edu | (872) 304-8591

Research Interests

Trustworthy Machine Learning: Machine Unlearning, Alignment & RLHF, Adversarial Machine Learning, Privacy

Education

Michigan State University (MSU) Aug. 2025 - Present
Incoming Ph.D. Student, Computer Science Advisor: Prof. Sijia Liu

Northwestern University (NU) Sep. 2023 - Jun. 2025
M.S., Computer Science Advisors: Prof. Qi Zhu and Prof. Xiao Wang

Tongji University Sep. 2019 - Jun. 2023
B.E., Software Engineering School of Computer Science and Technology

Research Experience

On the Adversarial Implications of Attention Sinks in LLMs Apr. 2025 - Present
Supervisor: Prof. Sijia Liu (MSU)

- Investigating attention sinks in LLMs to develop more effective backdoor poisoning attacks.
- Exploring applications in unlearned models where backdoor triggers can selectively recover forgotten knowledge.

Privacy-Preserving Tuning for Large Models Dec. 2023 - Mar. 2025
Supervisors: Prof. Qi Zhu (NU), Prof. Xiao Wang (NU)

- Developed Split Adaptation (SA) to ensure **data privacy** during adaptation of pre-trained Vision Transformers (ViTs), utilizing bi-level noise injection for privacy-preserving downstream tasks without data sharing.
- Protected **model privacy** by sharing only a low-bit quantized frontend of the ViT, preventing model leakage and ensuring secure adaptation.
- **Publication**: [1]

Professional Experience

Cloud Native Computing Foundation *Remote* Mar. 2023 - May 2023
Software Engineer Intern, Supervisor: Patrick Zheng
Project: KMS plugin for Notation CLI using Go.

SAP *Shanghai, China* Jun. 2022 - Mar. 2023
Cloud Developer Intern, Supervisor: April Qi
Project: Cloud Provider Exporter in Go on Kubernetes for AWS, Azure, and GCP, using Prometheus and Grafana.

Publications

*** indicates an equal contribution**

[1] Lixu Wang*, **Bingqi Shang***, Yi Li, Payal Mohapatra, Wei Dong, Xiao Wang, Qi Zhu. **Split Adaptation for Pre-trained Vision Transformers**. *CVPR'2025*.

Honors

- Shanghai Outstanding Graduate Award 2023
- Outstanding Undergraduate Dissertation Award of Tongji University 2023
- **National Scholarship** (Top 0.2%, highest undergraduate honor in China) 2020

Services

Journal Reviewer: IEEE TSP

Personal Interests

Astrophotography 2019 - Present

PROFESSIONAL SKILLS

Programming Languages: Python, Go, C++, Java, Rust, JavaScript, Latex, HTML, CSS
Machine Learning Systems: PyTorch, Transformers, W&B, OpenCV, Scikit-learn

Last updated: July 1, 2025.