

Bingqi Shang

 [Website](#) |  [Google Scholar](#) |  [GitHub](#) |  bshang@msu.edu

RESEARCH INTERESTS

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

Generative AI: Large-language Model, Multi-modal Language Model, Reasoning, Post-training

EDUCATION

Michigan State University (MSU)

Ph.D., Computer Science

Sep. 2025 - Present

Advisor: [Prof. Sijia Liu](#)

Northwestern University (NU)

M.S., Computer Science

Sep. 2023 - Jun. 2025

Advisors: [Prof. Qi Zhu](#) and [Prof. Xiao Wang](#)

Tongji University

B.E., Software Engineering

Sep. 2019 - Jun. 2023

[School of Computer Science and Technology](#)

RESEARCH EXPERIENCE

Backdooring Machine Unlearning via Attention Sinks

Apr. 2025 - Oct. 2025

Supervisor: [Prof. Sijia Liu](#) (MSU)

- Exposed a vulnerability in LLM unlearning: attention-sink-guided backdoors can **reactivate** forgotten behaviors.
- Exploring applications in unlearned models where backdoor triggers can selectively recover forgotten knowledge.
- **Publication:** [1].

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:** [2].

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] [Bingqi Shang*](#), [Yiwei Chen*](#), [Yihua Zhang](#), [Bingquan Shen](#), [Sijia Liu](#). [Forgetting to Forget: Attention Sink as A Gateway for Backdooring LLM Unlearning](#). *arXiv'2025*.

[2] [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 Transactions on Signal Processing

PERSONAL INTERESTS

Astrophotography

2019 - Present

PROFESSIONAL SKILLS

Programming Languages: Python, Go, C++, Java, Rust, \LaTeX , HTML, CSS, JavaScript

Machine Learning Systems: PyTorch, Transformers, Huggingface, W&B, OpenCV, Scikit-learn