# Mingjie Sun

Carnegie Mellon University
Pittsburgh, PA, 15213 USA
eric-mingjie.github.io
mingjies@andrew.cmu.edu
(+1) 412-652-2302

## EDUCATION

**Carnegie Mellon University** 2019 – Present
Ph.D. in Computer Science
Advisor: Prof. Zico Kolter
Expected Graduation: May 2025
Committee: Prof. Graham Neubig, Prof. Aditi Raghunathan, Prof. Kaiming He

**Tsinghua University** 2015 – 2019
Bachelor of Science in Computer Science
Yao Class, CS program led by Turing Award Laureate Prof. Andrew Yao

**University of Carlifornia, Berkeley** 2018
Exchange Student
Advisor: Prof. Trevor Darrell

## EXPERIENCE

**Robert Bosch**, Pittsburgh, PA, USA 06/2022 – 09/2022
Research Intern

**Microsoft Research**, Redmond, WA, USA 05/2021 – 08/2021
Research Intern

**Intel Research Labs**, Beijing, China 09/2018 – 05/2019
Research Intern

**University of California, Berkeley**, Berkeley, CA, USA 05/2018 – 08/2018
Research Visiting Student

## RESEARCH INTERESTS

My research interests lie in the area of deep learning. I am particularly interested in studying intriguing properties of deep neural networks and leveraging these insights to improve the performance of machine learning systems. Currently I am focused on understanding the challenges of compressing Large Language Models (LLMs).

## PUBLICATIONS

*** Equal contribution**

[1] **Mingjie Sun**, Xinlei Chen, Zico Kolter, Zhuang Liu.
Massive Activations in Large Language Models.
*First Conference on Language Modeling* (COLM), 2024.

[2] **Mingjie Sun***, Zhuang Liu*, Anna Bair, Zico Kolter.
A Simple and Effective Pruning Approach for Large Language Models.
*International Conference on Learning Representations* (ICLR), 2024.

[3] **Mingjie Sun**, Zico Kolter.
Single Image Backdoor Inversion via Robust Smoothed Classifiers.
*Conference on Computer Vision and Pattern Recognition* (CVPR), 2023.

[4] Nicholas Carlini*, Florian Tramer*, Krishnamurthy (Dj) Dvijotham, Leslie Rice, **Mingjie Sun**, Zico Kolter.
(Certified!!) Adversarial Robustness for Free!
*International Conference on Learning Representations* (ICLR), 2023.

[5] Sachin Goyal*, **Mingjie Sun***, Aditi Raghunathan, Zico Kolter.
Test-Time Adaptation via Conjugate Pseudo-labels.
*Neural Information Processing Systems* (NeurIPS), 2022.

[6] Xinlei Pan, Chaowei Xiao, Warren He, Shuang Yang, Jian Peng, **Mingjie Sun**, Jinfeng Yi, Zijiang Yang, Mingyan Liu, Bo Li, Dawn Song.
Characterizing Attacks on Deep Reinforcement Learning.
*International Conference on Autonomous Agents and Multiagent Systems* (AAMAS), 2022.

[7] **Mingjie Sun***, Zichao Li*, Chaowei Xiao*, Haonan Qiu, Bhavya Kailkhura, Mingyan Liu, Bo Li.
Can Shape Structure Features Improve Model Robustness under Diverse Adversarial Settings?
*International Conference on Computer Vision Conference* (ICCV), 2021.

[8] Hadi Salman, **Mingjie Sun**, Greg Yang, Ashish Kapoor, Zico Kolter.
Denoised Smoothing: A Provable Defense for Pretrained Classifiers.
*Neural Information Processing Systems* (NeurIPS), 2020.

[9] Zhuang Liu*, **Mingjie Sun***, Tinghui Zhou, Gao Huang, Trevor Darrell.
Rethinking the Value of Network Pruning.
*International Conference on Learning Representations* (ICLR), 2019.
(**Best Paper Award** at NIPS 2018 Workshop on Compact Deep Neural Networks with industrial applications.)

## TECHNICAL REPORTS

[1] Liqun Ma, **Mingjie Sun**, Zhiqiang Shen.
FBI-LLM: Scaling Up Fully Binarized LLMs from Scratch via Autoregressive Distillation.
*arXiv:2407.07093*, 2024.

[2] Rocktim Jyoti Das, **Mingjie Sun**, Liqun Ma, Zhiqiang Shen.
Beyond Size: How Gradients Shape Pruning Decisions in Large Language Models.
*arXiv:2311.04902*, 2023.

[3] Eungyeup Kim, **Mingjie Sun**, Aditi Raghunathan, Zico Kolter.
Reliable Test-Time Adaptation via Agreement-on-the-Line.
*arXiv:2310.04941*, 2023.

[4] **Mingjie Sun**, Siddhant Agarwal, Zico Kolter.
Poisoned classifiers are not only backdoored, they are fundamentally broken.
*arXiv:2010.09080*, 2020.

[5] **Mingjie Sun**, Jian Tang, Huichen Li, Bo Li, Chaowei Xiao, Yao Chen, Dawn Song.
Data Poisoning Attack against Unsupervised Node Embedding Methods.
*arXiv:1810.12881*, 2018.

## PATENTS

[1] **Mingjie Sun**, Sachin Goyal, Aditi Raghunathan, Jeremy Kolter, Wan-Yi Lin
System and Method for Test-time Adaptation via Conjugate Pseudolabels.
*US Patent 2024/0037416 A1*, issued, Feb. 1, 2024.

[2] **Mingjie Sun**, Jeremy Kolter, Filipe J Cabrita Condessa
Method and System for Breaking Backdoored Classifiers through Adversarial Examples.
*US Patent 2022/0100850 A1*, issued, Mar. 31, 2022.

## INVITED TALKS

- Understanding and Leveraging the Activation Landscape in Transformers
  Morgan Stanley Machine Learning Seminar Series 07/2024
  CMU, Computer Science Department, Proposal talk 05/2024
- Massive Activations in Large Language Models
  CMU Artificial Intelligence Seminar Series 04/2024
- A Simple and Effective Pruning Approach for Large Language Models
  Deep Learning: Classics and Trends (DLCT) 11/2023

## SERVICES

### Conference Reviewer

- Conference on Computer Vision and Pattern Recognition (CVPR) 2023, 2024
- International Conference on Learning Representations (ICLR) 2019, 2020, 2021, 2023, 2024
- Neural Information Processing Systems (NeurIPS) 2021, 2023, 2024
- European Conference on Computer Vision (ECCV) 2024
- International Conference on Computer Vision Conference (ICCV) 2023

### Journal Reviewer

- Transactions on Machine Learning Research (TMLR)
- IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)
- IEEE Transactions on Neural Networks and Learning Systems (TNNLS)

### Workshop Reviewer

- ICML Workshop on Efficient Systems for Foundation Models 2024
- NeurIPS Workshop on R0-FoMo 2023
- ICML Workshop on Deployable Generative AI 2023
- ICML Workshop on Principles of Distribution Shift (PODS) 2022
- NeurIPS Workshop on Compact DNNs with Industrial Applications 2018

## TEACHING

### Teaching Assistant

- Graduate Artificial Intelligence, 15-780, CMU. Spring 2024
- Deep Learning Systems, 10-414/714, CMU. Fall 2023