Calvin **Xu**

cx23@illinois.edu
github://cmxu | cmxu.io | stack://cmxu

Research Statement

I am focused on making trustworthy ML techniques ***usable*** in the ***real world***. My research lies at the intersection between ***ML*** and ***formal methods***. I am broadly interested in the generation of practical adversarial examples, certified training, and NN verification for the vision, language, and wireless.

Education

Current Sep 2021	PhD Student at University of Illinois at Urbana-Champaign Working on **PhD** in **CS**, GPA: 4.0/4.0
Dec 2019 Aug 2019	Advanced Studies Student at Massachusetts Institute of Technology Coursework - Advances in Computer Vision (6.869), GPA: 5.0/5.0
Dec 2018 Aug 2015	Undergraduate/Graduate Student at Washington University in St. Louis **MS** in **CS**, Certificate in Data Mining & Machine Learning, GPA: 3.9/4.0 **BS** in **Mathematics** and **BS** in **CS**, GPA: 3.8/4.0

Publications

Support is all you need for Certified VAE Training
Changming Xu, Debangshu Banerjee, Deepak Vasisht, Gagandeep Singh. ICLR '25

Certified DNN Training against Universal Adversarial Perturbations
Changming Xu, Gagandeep Singh. ECCV '24

Scalable Relational Verification and Training for Deep Neural Networks
Debangshu Banerjee, Changming Xu, Gagandeep Singh. SAIV@CAV '24

Robust Universal Adversarial Perturbations
Changming Xu, Gagandeep Singh. ICML '24

Input-Relational Verification of Deep Neural Networks
Debangshu Banerjee, Changming Xu, Gagandeep Singh. PLDI '24

Bypassing the Safety Training of Open-Source LLMs with Priming Attacks
Jason Vega*, Isha Chaudhary*, Changming Xu*, Gagandeep Singh. ICLR Tiny Paper '24 (invite to present)

Exploring Practical Vulnerabilities of Machine Learning-based Wireless Systems
Zikun Liu, Changming Xu, Emerson Sie, Deepak Vasisht, Gagandeep Singh. NSDI '23

Race Detection and Reachability in Nearly Series-Parallel DAGs
Kunal Agrawal, Joseph Devietti, Jeremy T. Fineman, I-Ting Angelina Lee, Robert Utterback, Changming Xu. SODA '18

*equal contribution

Fellowships & Awards

Aug 2022	Qualcomm Innovation Fellowship Winner (one of 19 teams out of ∼300) ***Proposal***: Provably Robust Machine Learning for Wireless Systems
Aug 2019	1st place at AdvML Challenge at ***SigKDD 2019***
Apr 2018	Dean's Select Fellowship for Research Excellence (WUSTL)
Dec 2015-17	Putnam Exam: 28, 20, 12
Aug 2015	Compton Scholar for Mathematics and Physics (4 per grade)
Apr 2017	Missouri Math Competition: 1st Place Team
Dec 2015	ICPC Regional: Top 5 Team

Research Experience

Current – Aug 2021 | PhD Student at University of Illinois in Urbana-Champaign

Currently working on

- Certified training and efficient DNN verification of universal adversarial perturbations, VAEs, and other attacks/architectures
- Training for better LLM alignment
- Training networks that are certifiably robust under network compression (pruning/quantization)
- Probabilistic verification of VAEs in the wireless domain
- Certifiably robust training of DNNs

Mar 2018 – Jun 2017 | Research Assistant at Carnegie Mellon University

- Employed Adversarial ML techniques to thwart defect prediction
- Developed theory for attacking and defending high dimensional SVMs
- Implemented data poisoning attacks on the Drebin Android Malware dataset, and presented poster at CMU
- Proved that ~10 poisoned data points can be enough to significantly reduce the effectiveness of the malware detector

May 2017 – Jun 2016 | Research Assistant at Washington University in St. Louis

- Created benchmarks and proved correctness for a work-stealing scheduler, improving cache efficiency by factor of 2
- Rigorously proved a nearly series parallel race detection algorithm which matched asymptotic runtime of series parallel case
- Published paper in *ACM-SIAM SODA '18*

Work Experience

Dec 2024 – Sept 2024 | Intern at Bytedance

- Finetuning and soft prompting VLMs for content moderation

May 2021 – Mar 2019 | Associate Staff at MIT Lincoln Laboratory

- Applied DAGAN to network traffic to augment unbalanced classes resulting in a ~20% increase in classification accuracy
- Applied CNNs, Word2Vec, and Q-Learning techniques to flight data to predict flight reroutes, published at *INFORMS '19*
- Gathered dataset and built system using NLP and Neural Networks (NNs) to determine when to apply Adversarial ML
- Explored Semantic Adversarial Examples and Spatial Transformer Networks for data augmentation
- Developed an algorithm using GANs and Non-Negative NNs that won 1st place at AdvML Challenge at *SigKDD 2019*

Dec 2018 – Sep 2016 | Teaching Assistant at Washington University in St. Louis

- Graded and held office hours for graduate level Machine Learning (150 students) and Numerical Applied Mathematics (50 students)
- Designed, ran, and graded the final project for the Machine Learning course

Aug 2018 – May 2018 | Intern at EUB-INC in Beijing, China

- Leveraged NNs and clustering to automate user grouping for data-driven advertisement on WeChat.
- Developed algorithm which reduced turnaround time by 90% for data team, allowing for faster, more targeted advertisement

Skills

Languages | English (Native), Chinese (Fluent), Japanese (Basic Knowledge)