Last updated in December 2024

Yihan Wang

📍 Waterloo, Canada (for now) ✉ yihanwang@amss.ac.cn 🏠 yihan-wang.com EhanW

Education

Chinese Academy of Sciences *Sept. 2019 – Present*
PhD in Applied Mathematics
Advisor: Prof. Xiao-Shan Gao

University of Waterloo *Dec. 2023 – Dec. 2024*
Visiting Student
Host Advisor: Prof. Yaoliang Yu

Peking University *Sept. 2017 – Feb. 2018*
Visiting Student

Sichuan University *Sept. 2015 – June 2019*
BS in Mathematics and Applied Mathematics

Publications

Efficient Availability Attacks against Supervised and Contrastive Learning Simultaneously 📎
Yihan Wang, Yifan Zhu, Xiao-Shan Gao
Proceedings of the 38th Conference on Neural Information Processing Systems (NeurIPS 2024)

Data-Dependent Stability Analysis of Adversarial Training 📎
Yihan Wang, Shuang Liu, Xiao-Shan Gao
Neural Networks

Machine Unlearning for Contrastive Learning under Auditing 📎
Yihan Wang*, Yiwei Lu*, Guojun Zhang, Franziska Boenisch, Adam Dziedzic, Yaoliang Yu, Xiao-Shan Gao
ICML 2024 Next Generation of AI Safety Workshop (Oral)

On the Robustness of Neural Networks Quantization against Data Poisoning Attacks 📎
Yiwei Lu, ***Yihan Wang***, Guojun Zhang, Yaoliang Yu
ICML 2024 Next Generation of AI Safety Workshop

Game-Theoretic Unlearnable Example Generator 📎
Shuang Liu, ***Yihan Wang***, Xiao-Shan Gao
Proceedings of the 38th Annual AAAI Conference on Artificial Intelligence (AAAI 2024)

Adversarial Parameter Attack on Deep Neural Networks 📎
Lijia Yu, ***Yihan Wang***, Xiao-Shan Gao
Proceedings of the 40th International Conference on Machine Learning (ICML 2023)

Restore Translation Using Equivariant Neural Networks 📎
Yihan Wang, Lijia Yu, Xiao-Shan Gao
Proceedings of the 30th International Conference on Neural Information Processing (ICONIP 2023)

Mitigating Robust Overfitting in Wasserstein Distributionally Robust Optimization 📎
Shuang Liu, ***Yihan Wang***, Xiao-Shan Gao
Preprint

Projects

Alignment Calibration

- Developed an unlearning algorithm for contrastive learning that is easy to audit for data owners.

Augmented Unlearnable Examples & Augmented Adversarial Poisoning

- Developed two effective and efficient availability attacks against supervised and contrastive learning.

Adversarial Parameter Attack

- Developed an algorithm to reduce the robustness of a model while maintaining accuracy.

Award and Honors

Loo-Keng Hua Scholarship from AMSS, CAS	*2020 – 2024*
Top-Notch Scholarship from Sichuan University	*2017 – 2019*
Comprehensive First-class Scholarship from Sichuan University	*2016*

Professional Service

I regularly served as a reviewer for International Conference on Machine Learning (ICML), International Conference on Learning Representation (ICLR), Neural Information Processing Systems (NeurIPS).