Mintong Kang

✉ mintong2@illinois.edu | 🌐 personal website | kangmintong | ☎ (+1)217-979-7287

Education

University of Illinois at Urbana-Champaign August 2022 - present

Computer Science Ph.D.

- Advised by Prof. **Bo Li**
- Research Interest: the intersection of machine learning and robustness, fairness, generalization
- Affiliated at Department of Computer Science

Zhejiang University August 2018 - June 2022

Bachelor of Computer Science and Technology

- Advised by Prof. **Xi Li**
- GPA: 3.95/4.0, 91.5/100
- Affiliated at Department of Computer Science and Chu Kochen Honors College
- Outstanding Undergraduate Award and Outstanding Thesis Award at Zhejiang University

Publications and Preprints

COLEP: Certifiably Robust Learning-Reasoning Conformal Prediction via Probabilistic Circuits
Mintong Kang, Nezihe Merve Gürel, Linyi Li, Bo Li
[**ICLR 2024**] (Twelfth International Conference on Learning Representations)

DecodingTrust: A Comprehensive Assessment of Trustworthiness in GPT Models
Boxin Wang*, Weixin Chen*, Hengzhi Pei*, Chulin Xie*, **Mintong Kang***, Chenhui Zhang*, Chejian Xu, Zidi Xiong, Ritik Dutta, Rylan Schaeffer, Sang T. Truong, Simran Arora, Mantas Mazeika, Dan Hendrycks, Zinan Lin, Yu Cheng, Sanmi Koyejo, Dawn Song, Bo Li
[**NeurIPS 2023**] **(Outstanding Paper Award)** (Thirty-seventh Conference on Neural Information Processing Systems)

DiffAttack: Evasion Attacks Against Diffusion-Based Adversarial Purification
Mintong Kang, Dawn Song, Bo Li
[**NeurIPS 2023**] (Thirty-seventh Conference on Neural Information Processing Systems)

Certifying Some Distributional Fairness with Subpopulation Decomposition
Mintong Kang*, Linyi Li*, Maurice Weber, Yang Liu, Ce Zhang, Bo Li
[**NeurIPS 2022**] **(Spotlight)** (Thirty-sixth Conference on Neural Information Processing Systems)

Fairness in Federated Learning via Core-Stability
Bhaskar Ray Chaudhury, Linyi Li, **Mintong Kang**, Bo Li, Ruta Mehta
[**NeurIPS 2022**] **(Spotlight)** (Thirty-sixth Conference on Neural Information Processing Systems)

C-RAG: Certified Generation Risks for Retrieval-Augmented Language Models
Mintong Kang, Nezihe Merve Gürel, Ning Yu, Dawn Song, Bo Li
In preprint, 2024.

Certifiably Byzantine-Robust Federated Conformal Prediction
Mintong Kang, Zhen Lin, Jimeng Sun, Cao Xiao, Bo Li
In preprint, 2023.

FaShapley: Fast and Approximated Shapley Based Model Pruning Towards Certifiably Robust DNNs
Mintong Kang, Linyi Li, Bo Li
[**SaTML 2023**] (IEEE Conference on Secure and Trustworthy Machine Learning 2023)

Data, Assemble: Leveraging Multiple Datasets with Heterogeneous and Partial Labels
Mintong Kang, Yongyi Lu, Alan L. Yuille, Zongwei Zhou
[**ISBI 2023**] (IEEE International Symposium on Biomedical Imaging 2023)

MgSvF: Multi-Grained Slow vs. Fast Framework for Few-Shot Class-Incremental Learning
Hanbin Zhao, Yongjian Fu, **Mintong Kang**, Qi Tian, Fei Wu, Xi Li
[**TPAMI 2021**] (IEEE Transactions on Pattern Analysis and Machine Intelligence 2021)

*Note: * stands for equal contribution*

Internship Experience

Relativity, Chicago **Remote**

Research Intern advised by Prof. ***Cao Xiao*** May 2023 - August 2023

- Research on trustworthy federated learning
- Accomplished the paper *Certifiably Byzantine-Robust Federated Conformal Prediction*
- The first certifiably robust federated conformal prediction framework (Rob-FCP) in the Byzantine setting
- Propose a maliciousness score to identify Byzantine clients in federated conformal prediction
- Theoretically provide the coverage guarantees of Rob-FCP in both IID and non-IID settings
- Empirically demonstrate the robustness of Rob-FCP in federated Byzantine settings

AIsecure Lab, University of Illinois at Urbana-Champaign **Remote**

Research Intern advised by Prof. ***Bo Li*** November 2021 - May 2022

- *Two* projects accomplished: certified fairness and certifiably robust pruning
- *Certified fairness*: Provide the first end-to-end fairness certification framework
- Propose the subpopulation decomposition method to solve the non-convex optimization
- Tight certificate in realistic scenarios (six practical datasets)
- *Certifiably robust pruning* Propose a novel pruning criterion to achieve high certified robustness
- Propose sample-size optimization and gradient-based estimation
- Achieve SOTA for certifiably robust pruning in multiple settings

CCVL Lab, Johns Hopkins University **Remote**

Research Intern advised by Prof. ***Alan L. Yuille*** *and Dr.* ***Zongwei Zhou*** May 2021 - September 2021

- Target the partial label problem in medical imaging datasets
- Propose the initiative to assemble partially labeled data from multiple sources
- Propose a framework to effectively learn from assembled heterogenous data
- Exhaustive demonstration of the initiative for classification, segmentation, and detection
- Achieve SOTA for multi-label classification on Chest X-ray dataset

DCD Lab, Zhejiang University **Hangzhou, China**

Research Intern advised by Prof. ***Xi Li*** September 2020 - February 2021

- Target the incremental learning problem especially the regularization-based approach
- Propose a novel perspective to view the regularization from the signal frequency dimension
- Propose effective regularization in the frequency domain
- Outperform baselines with regularization in the feature space by a large margin

Selected Awards

- NeurIPS 2023 Outstanding Paper Award
- NeurIPS 2023 Scholar Award
- NeurIPS 2022 Scholar Award
- Outstanding Undergraduate Award at Zhejiang University
- Scholarship of Zhejiang Province Government
- First Prize Scholarship at Zhejiang University
- First Prize in Chinese Olympic Mathematics Competition

Service

- **Conference Reviewers**: ICML 2022-2023; NeurIPS 2022-2023; ICLR 2024; KDD 2023
- **Workshop organize**: KLR@ICML 2023

Language and Skills

- TOEFL iBT: Total 111/120, Reading 30/30, Listening 29/30, Speaking 24/30, Writing 28/30
- GRE General Test: Total 327/340, Verbal 157/170, Quantitative 170/170
- Programming skills: Python, Pytorch, Tensorflow, Keras, C/C++

Last updated: February 5, 2024