

Shaopeng Fu

Last Update: December 2024

🏠 shaopengfu.me
🔑 i7cWm4gAAAAJ
🐙 github.com/fshp971

✉️ shaopeng.fu@kaust.edu.sa
✉️ shaopengfu15@gmail.com
☎️ +966 (0) 56 534 0337

EDUCATION

King Abdullah University of Science and Technology

Provable Responsible AI and Data Analytics (PRADA) Lab

Ph.D. Student in Computer Science

Advisor: Prof. Di Wang

Thuwal, Saudi Arabia

Aug. 2023 – Present

The University of Sydney

UBTECH Sydney Artificial Intelligence Centre

Master of Philosophy (Engineering and IT)

Advisor: Prof. Dacheng Tao

Thesis: Bayesian Inference Forgetting

Sydney, Australia

Oct. 2019 – Jan. 2021

South China University of Technology

B.Sc in Mathematics and Applied Mathematics

Advisor: Prof. Chuhua Xian (Advising the Competitive Programming Group affiliated to School of CSE)

GPA: 3.61/4.00 | Rank: 6/46

Guangzhou, China

Sep. 2015 – Jun. 2019

WORK EXPERIENCES

King Abdullah University of Science and Technology

Research Intern (Topic: Adversarial Robustness; Advisor: Prof. Di Wang)

Thuwal, Saudi Arabia

May 2023 – Aug. 2023

JD.com, Inc.

Algorithm Engineer @ JD Explore Academy (Full-time)

Beijing, China

Mar. 2021 – Jul. 2022

- First-author of two ICLR 2022 papers.
- Co-author of the *White Paper on Trustworthy Artificial Intelligence* ([Chn Ver.](#)) ([Eng Ver.](#)) (with CAICT).
- Chief developer of **TAICore**, a trustworthy AI assessment toolkit powered by JD Explore Academy for assessing the robustness and privacy-preserving ability of white-box and black-box ML models.

The University of Sydney

Research Assistant (Topic: Machine Unlearning; Advisor: Prof. Dacheng Tao)

Sydney, Australia

Oct. 2019 - Oct. 2020

RESEARCH INTERESTS

My research lies in trustworthy AI. I am interested in using mathematical principles to identify and mitigate security and privacy risks in real-world machine learning systems. Currently, I am working on:

- Adversarial Robustness of Pre-trained Large Models (LLMs/MLLMs)
- Privacy-preserving Ability of Pre-trained Large Models (LLMs/MLLMs)

PUBLICATIONS

CONFERENCES & JOURNALS

1. **Shaopeng Fu** and Di Wang. Theoretical Analysis of Robust Overfitting for Wide DNNs: An NTK Approach. In *International Conference on Learning Representation (ICLR)*, 2024.
2. **Shaopeng Fu**, Fengxiang He, Yang Liu, Li Shen, and Dacheng Tao. Robust Unlearnable Examples: Protecting Data Against Adversarial Learning. In *International Conference on Learning Representation (ICLR)*, 2022.

3. **Shaopeng Fu***, Fengxiang He*, and Dacheng Tao. Knowledge Removal in Sampling-based Bayesian Inference. In *International Conference on Learning Representation (ICLR)*, 2022.
4. Zeke Xie, Fengxiang He, **Shaopeng Fu**, Issei Sato, Dacheng Tao, and Masashi Sugiyama. Artificial Neural Variability for Deep Learning: On Overfitting, Noise Memorization, and Catastrophic Forgetting. *Neural Computation* 33 (8), 2021.

MANUSCRIPTS

1. **Shaopeng Fu**, Xuexue Sun, Ke Qing, Tianhang Zheng, and Di Wang. Pre-trained Encoder Inference: Revealing Upstream Encoders In Downstream Machine Learning Services. *arXiv preprint arXiv:2408.02814*, 2024.
2. Fengxiang He*, **Shaopeng Fu***, Bohan Wang*, and Dacheng Tao. Robustness, Privacy, and Generalization of Adversarial Training. *arXiv preprint arXiv:2012.13573*, 2020.

SELECTED AWARDS

International Collegiate Programming Contest (ICPC)

- **Silver Medal**, The ICPC Asia-East Continent Final Xi'an Site Dec. 2018
- **Silver Medal**, The ICPC Asia Regional Contest Qingdao Site Nov. 2018
- **Gold Medal (Rank: 6/186)**, The ICPC Asia Regional Contest Shenyang Site Oct. 2018
- **Silver Medal**, The ACM-ICPC Asia Regional Contest Xi'an Site Oct. 2017

2017-2018 China National Scholarship Ministry of Education of P.R. China, Nov. 2018

2016-2017 China National Scholarship Ministry of Education of P.R. China, Nov. 2017

SERVICES

Conference Reviewer

- ICML (2022, 2023, 2024, 2025), ICLR (2022, 2023, 2024, 2025), NeurIPS (2021, 2022, 2023, 2024), AIS-TATS (2021, 2024, 2025).

Conference Committee Member

- ACM CCS (2024 Artifact Evaluation), AAAI (2025).

Journal Reviewer

- IEEE TPAMI (2024), IEEE TCYB (2021), Springer NPL (2020).

TEACHING

Teaching Assistant of CS 229: Machine Learning, Spring 2024 @ KAUST

MISCELLANEOUS

Competitive Programming: My [Codeforces](#) account is [fshp971](#).

Programming Languages: C/C++ (Mainly for Competitive Programming), Python (Mainly for AI Research).

Others: PyTorch, JAX, Vim, Linux, Arch Linux.