WENJIE WANG

(021)-20685694 ⋄ wangwj1@shanghaitech.edu.cn
1C-503E, School of Information Science and Technology
ShanghaiTech University, Shanghai, China

EDUCATION

Emory University,Atlanta, GA, USA *Aug 2017 - Nov 2022*
Ph.D in Computer Science
Huazhong University of Science and Technology (HUST)*, Wuhan, China* Sep 2013 - Jun 2017
B.S in Biotechnology
University of California, Berkeley(UCB)*, Berkeley, CA, USA* Aug 2015 - Dec 2015
Department of Integrative Biology

WORK EXPERIENCE

Shanghai Tech University, Shanghai Feb 2023 – Present
Assistant Professor / Researcher

Tencent AI Lab, Seattle Jun 2020 – August 2020
Research Intern in Speech Group *Manager: Dong Yu*

IBM T.J Watson Research Center Jun 2019 – August 2019
Research Intern in Security AI group *Manager: Ian Molloy*

RESEARCH FOCUS

AI Security

- Adversarial Machine learning, Adversarial attacks, Certified Robustness,

AI Privacy

- Differential Privacy, Federated Learning, Machine Unlearning

PUBLICATIONS

- "*LinkPrompt: Natural and Universal Adversarial Attacks on Prompt-based Language Models*", Xu Yue, **Wenjie Wang**†, North American Chapter of the Association for Computational Linguistics (NAACL) 2024
- "*IGAMT: Privacy-Preserving Electronic Health Record Synthesization with Heterogeneity and Irregularity*", **Wenjie Wang**†, Pengfei Tang, Jian Lou, Yuanming Shao, Lance Waller, Yi-An Ko, Li Xiong, Association for the Advancement of Artificial Intelligence (AAAI) 2024
- "*Demo: Certified Robustness on Toolformer*", Xu Yue, **Wenjie Wang**†, ACM SIGSAC Conference on Computer and Communications Security (CCS) 2023 Demo/Poster Session
- "*Efficient Text Analysis with Pre-trained Neural Network Models*", Jia Cui, Heng Lu, **Wenjie Wang**, Shiyin Kang, Liqiang He, Guangzhi Li, Dong Yu, IEEE Spoken Language Technology Workshop (SLT) 2022
- "*Certified Robustness to Word Substitution Attack with Differential Privacy*", **Wenjie Wang**, Pengfei Tang, Jian Lou, Li Xiong, North American Chapter of the Association for Computational Linguistics (NAACL) 2021
- "*Generating Adversarial Examples with Distance Constrained Adversarial Imitation Networks*", Pengfei Tang, **Wenjie Wang**, Li Xiong, IEEE Transactions on Dependable and Secure Computing (TDSC Journal) 2021
- "*RADAR: Recurrent Autoencoder based Detector for Adversarial Temporal EHR*", **Wenjie Wang**, Pengfei Tang, Li Xiong, Xiaoqian Jiang, The European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML)2020
- "*Utilizing Multimodal Model Consistency to Detect Adversarial Examples*", **Wenjie Wang**, Youngja Park, Taesung Lee, Ian Molloy, Pengfei Tang, Li Xiong, Empirical Methods in Natural Language Processing(EMNLP) workshop: ClinicalNLP2020

- "*Bacterial contamination screening and interpretation for biological laboratory environments*", Xue Zhu, Xi Li, **Wenjie Wang**, Kang Ning, Medicine in Microecology 2020

PRE-PRINTS

- "*Certified PEFTSmoothing: Parameter-Efficient Fine-Tuning with Randomized Smoothing*", Chengyan Fu, **Wenjie Wang**†, arXive
- "*Don't Say No: Jailbreaking LLM by Suppressing Refusal*", Yukai Zhou, **Wenjie Wang**†, arXive
- "*Wasserstein Adversarial Examples on Univarient time series data*", **Wenjie Wang**, Jian Lou, Pengfei Tang, Li Xiong, arXive
- "*Killing Two Birds with One Stone: Achieving both Differential Privacy and Certified Robustness via Input Perturbation*", Pengfei Tang*, **Wenjie Wang***, Xiaolan Gu, Jian Lou, Li Xiong, Ming Li, under submission

TEACHING

2023Fall: CS246 Trustworthy Machine Learning Graduate Course

- Number of Students: 33, Course Score: 4.76, Teacher Score: 4.83

2024Spring: CS150A Database Undergraduate Course

- Number of Students: 65

SERVICES

Jiangsu undergraduate enrollment propaganda group

- Server as the speaker 5 times

Hubei graduate enrollment propaganda group

- Serve as the speaker 1 time

Conference Program Committee Member

- The Association for Computational Linguistics (ACL2024)
- North American Chapter of the Association for Computational Linguistics (NAACL2024)
- The Association for the Advancement of Artificial Intelligence (AAAI2023,2024)
- The Conference on Information and Knowledge Management (CIKM2022)
- The European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML2020,2023)
- American Medical Informatics Association (AMIA2020 2021)

*Equally Contribution
†Corresponding Author