Arezoo Rajabi

Senior Quantitative Analytic Specialist AI/ML & Adjunct faculty at University of Washington

Email: rajabia@uw.edu
Linkedin: www.linkedin.com/in/arezoo-rajabi
Homepage: http://rajabia.github.io

Wells Fargo Bank
333 Market ST
San Francisco, CA

EDUCATION

Oregon State University, Corvallis, Oregon, USA Sep. 2014 – June 2021
Ph.D. in Computer Science
Thesis: Two Sides of a Coin: Adversarial-Based Image Privacy and Defending Against Adversarial Perturbations for Robust CNNs

Sharif University of Technology, Tehran, Iran Sep. 2011 – Sep. 2013
M.Sc. in Computer Engineering (Software Engineering)
Thesis: Local Community Detection in Social Networks

Sharif University of Technology, Tehran, Iran Sep. 2005 – Jan. 2011
B.Sc. in Computer Science

RESEARCH AREAS

Attacks and Defenses in Deep Learning, Large Language Models,Differential Privacy

WORK& RESEARCH EXPERIENCE

Senior Quantitative Data Analytic Specialist AI/ML, Dec. 2022 – Peresent
Wells Fargo Bank, CA, USA

- *Machine Learning Model Development:* Developing and deploying machine learning models for privacy-sensitive and large datasets.
- *Performance Monitoring:* Designing comprehensive monitoring plans to assess the performance of deployed models.
- *Model Lifecycle Management:* Documenting the complete model lifecycle, including design solutions and key performance indicators (KPIs).

Postdoctoral Scholar, March 2021 – Dec. 2022
NSL Lab, University of Washington, Seattle, WA, USA

- *Multi-Domain Trojan Detection:* Proposing a multi-domain Trojan sample detection system during the inference phase. Achieving a minimum success rate of 85% for detecting Trojan samples in text, images, and audio domains.
- *Privacy-Preserving RL Algorithm:* Developing a differential privacy method for RL algorithms with a risk-neutral decision-making approach and creating a defense mechanism against membership inference attacks for pre-trained DNNs.
- *Federated Learning for Trojan Prevention:* Implementing a federated learning approach to counteract Trojan samples during the training phase.

Graduate Research Assistant, Sep. 2014 – Sep. 2020
Oregon State University, Corvallis, Oregon, USA

- Developing image privacy methods based on adversarial learning methods against automated face detection methods
- Developing two fault tolerance approaches for outliers in distributed smart grid power systems

SKILLS

Domain Specific Skill: Image Classification, Large Language Models, Statistical Analysis and Testing, Clustering and Anomaly Detection, Graph Convolutional Networks, Reinforcement Learning
Programming Languages: Python, Java, R, Matlab, C#
Machine/Deep Learning Tools: PyTorch, Opacus, Keras, Tensorflow, NeMo, ggplot, SciPy, Robustness, LangChain, Hugging Face, Colab
Other Tools: SQL, Hadoop, Amazon Web Services, GCP, H2O , Jira (Agile Methodologies)
Soft Skills: Critical Thinking, Problem Solving

SELECTED PUBLICATIONS

1. A. Rajabi, S. Asokraj, F. Jiang, L. Niu, B. Ramasubramanian, J. Ritcey, R. Poovendran, MDTD: A Multi-Domain Trojan Detector for Deep Neural Networks, ACM Conference on Computer and Communications Security (ACM CCS), Sep. 2023.

2. J. Jia, Z. Yuan, D. Sahabandu, L. Niu, **A. Rajabi**, B. Ramasubramanian, B. Li, R. Poovendran, FLGAME: A Game-theoretic Defense against Backdoor Attacks In Federated Learning, Thirty-seventh Conference on Neural Information Processing Systems (NeurIPS), Sep 2023 (https://neurips.cc/virtual/2023/poster/70499).

3. A. Rajabi, D. Sahabandu, L. Niu, B. Ramasubramanian, R. Poovendran, LDL: A Defense for Label-Based Membership Inference Attacks, ACM Asia Conference on Computer and Communications Security (AsiaCCS), July 2023 (**7% acceptance rate**).

4. A. Rajabi, B. Ramasubramanian, A. Marruf, R. Poovendran, Privacy Preserving Reinforcement Learning Beyond Expectation, Accepted in 61st IEEE Conference on Decision and Control, 2022.(https://arxiv.org/pdf/2203.10165.pdf).

5. A. Rajabi, M. Abbasi, R. B. Bobba, K. Tajik, Adversarial Images Against Super-Resolution Convolutional Neural Networks for Free, Privacy Enhancing Technology Symposium (PETS), 2022.

6. M. Abbasi, **A. Rajabi**, C. Shui, C. Gagné, R. B. Bobba, Toward Adversarial Robustness by Diversity in an Ensemble of Specialized Deep Neural Networks, Canadian Conference on Artificial Intelligence (Canadian AI), 2020. (Best Paper Award)

PATENTS

Arezoo Rajabi, Dinuka Sahabandu, Luyao Niu, Bhaskar Ramasubramanian, Radha Poovendran, *LDL: A Defense for Label-Based Membership Inference Attacks*, Record of Innovation filed with CoMotion At University of Washington, Seattle Dec. 2022.

PROFESSIONAL SERVICES

Adjunct Faculty at University of Washington 2022- present Organizer at The Trojan Detection Challenge (LLM Edition), NeurIPS 2023 2023
Organizer at Trojan Detection Challenge, NeurIPS 2022 2022
Diversity co-chair at Security and Privacy Symposium 2023

AWARDS

First Place Winner at Graduate Research Showcase for Poster Presentation 2018
Cyber Resilient Energy Delivery Consortium (CREDC) Summer School Student Scholarship 2017
Student Travel Awards from Top Security Conferences (S&P, CCS, GREPSEC, and ACSAC)