Kasra Ahmadi

Ph.D. Candidate, Computer Science and Engineering - Machine Learning Security & Post-Quantum Cryptography

+1-8136-146240 | kasra.research@gmail.com | linkedin.com/in/kasra-ahmadii | github.com/KasraAhmadi | Google Scholar | kasraahmadi.github.io | Sunnyvale, CA

SUMMARY

CS Ph.D. candidate and ML research engineer focused on privacy-preserving and security-critical machine learning. First author on 10+ peer-reviewed publications, including a Best Paper Award at IEEE S&P 2025 Workshops. Builds production ML systems — PII detection classifiers, decoder-model probes, model routing for Claude Code — and researches how ML systems fail and get attacked: adversarial evasion of RAG retrieval, differentially private federated learning, and fault-injection/side-channel resilience for both DNN inference and post-quantum cryptographic hardware (Kyber, Dilithium, ECC). Targeting ML engineering, applied research, and ML security roles.

CORE COMPETENCIES

Machine Learning & LLM Systems • Adversarial ML & Evaluation • Privacy-Preserving ML (DP, Federated Learning) • PII Detection & Content Safety • ML Security & Fault/Side-Channel Analysis • Post-Quantum Cryptography • Production ML Engineering • Research & Publication (10+ papers)

EDUCATION

Ph.D. in Computer Science and Engineering - *University of South Florida* Jan. 2022 - Oct. 2026
Tampa, FL

Master of Information Technology Engineering - *Amirkabir University of Technology* Sep. 2018 - Aug. 2020
Tehran, Iran

RESEARCH AND INDUSTRY EXPERIENCE

RealmLabs *May 2026 - Present*

ML Research Engineer Intern - Sunnyvale, CA

- Developed PII privacy and security workflows using frozen Qwen 3.5 and Llama 3.1 models, achieving a 3% improvement on F1 score over ModernBERT by disabling causal masking to capture full bidirectional context.
- Developing probe models on top of decoder hidden states to detect PII at the token level.
- Working on cost optimization and model routing for Claude Code by classifying session intent and routing sessions to the appropriate Claude model.
- Integrating vLLM with sparse autoencoders (SAEs) to capture token-level signals during generation in a single pass.

University of South Florida *Jan. 2022 - Present*

Research Assistant - Tampa, FL

- Designed a human-in-the-loop, differentially private federated learning framework for fine-tuning LLMs (BERT) on memory-constrained devices, solving the fixed-memory limitation of existing DP mechanisms; **won Best Paper Award at IEEE S&P 2025 Workshops (HMI-SA)**.
- Built MAED, an algorithm-level runtime error-detection framework using mathematical identities to catch fault-injection attacks and hardware faults on DNN activation functions (ReLU, sigmoid, tanh), hardening embedded ML inference pipelines against adversarial and natural faults.
- Developed lightweight, algorithm-level side-channel and fault-attack countermeasures for NIST-standardized post-quantum cryptosystems (CRYSTALS-Kyber, Dilithium) and classical public-key schemes (ECC/ECSM), published in IEEE TVLSI, IEEE TCAD, and ACM TECS.
- Authored 10+ peer-reviewed papers spanning ML security, privacy-preserving AI, and hardware security under an NSF-funded research program; serve as a peer reviewer for IEEE TVLSI, TCAS-I, and ACM TECS.

TD Synnex *May 2025 - Aug. 2025*

AI Engineer Intern - Clearwater, FL

- Designed and implemented cloud-native AI workflows combining LLMs, RAG, and multi-agent orchestration.

AgWise *May 2024 - Aug. 2024*

Software Engineer Intern - St. Petersburg, FL

- Built an AWS-based ETL pipeline using Lambda, Glue, and S3 to process laboratory nutrient data for crop recommendation models.

PaarLift *Jan. 2018 - Apr. 2020*

ML Research Engineer - Tehran, Iran

- Developed AI-driven elevator scheduling algorithms that reduced passenger wait times by 27% across deployments in 100+ commercial buildings.

SELECTED RESEARCH PROJECTS

Federated LLM Fine-Tuning with Differential Privacy — *GitHub* | *Paper*

Built a human-in-the-loop framework for privacy-utility trade-off control in federated fine-tuning of BERT-based models using Flower, DP-SGD, and LoRA. Best Paper Award, IEEE S&P 2025 Workshops.

Gradient-Based Adversarial Attacks Against RAG Systems — *GitHub*

Designed evasion attacks that perturb vector embeddings to degrade RAG retrieval quality, answer reliability, and downstream system accuracy.

MAED: Error Detection for Fault Attacks on DNN Inference — *GitHub*

Developed algorithm-level error detection methods for activation-function faults in DNN inference for safety- and security-critical deployments.

PII-360 — *GitHub* | *Product*

Built an open-source Chrome extension that detects personally identifiable information in images and PDFs using ONNX models running locally on-device.

Fault-Attack Simulation for Classical and Post-Quantum Cryptography — *GitHub*

Simulated fault-injection attacks on Kyber, Dilithium, and ECC, then proposed lightweight algorithm-level countermeasures with near-100% error detection.

SELECTED PUBLICATIONS

1. **An Interactive Framework for Implementing Privacy-Preserving Federated Learning: Experiments on Large Language Models.** Kasra Ahmadi, Rouzbeh Behnia, Reza Ebrahimi, Mehran Mozaffari Kermani, Jeremiah Birrell, Jason Pacheco, Attila A. Yavuz. *IEEE Security and Privacy Workshops (SPW), 2025. Best Paper Award.* arXiv GitHub

2. **PUF-Kyber: Design of a PUF-Based Kyber Architecture Benchmarked on Diverse ARM Processors.** Saeed Aghapour, Kasra Ahmadi, Mila Anastasova, Mehran Mozaffari Kermani, Reza Azarderakhsh. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2024.* IEEE

3. **PUF-Dilithium: Design of a PUF-Based Dilithium Architecture Benchmarked on ARM Processors.** Saeed Aghapour, Kasra Ahmadi, Mehran Mozaffari Kermani, Mila Anastasova, Reza Azarderakhsh. *ACM Transactions on Embedded Computing Systems, vol. 24, no. 2.* ACM

4. **Efficient Error Detection Schemes for ECSM Window Method Benchmarked on FPGAs.** Kasra Ahmadi, Saeed Aghapour, Mehran Mozaffari Kermani, Reza Azarderakhsh. *IEEE Transactions on VLSI Systems, vol. 32, no. 3, pp. 592-596, March 2024.* IEEE GitHub

5. **Efficient Error Detection Cryptographic Architectures Benchmarked on FPGAs for Montgomery Ladder.** Kasra Ahmadi, Saeed Aghapour, Mehran Mozaffari Kermani, Reza Azarderakhsh. *IEEE Transactions on VLSI Systems, early access, 2024.* IEEE GitHub

6. **Efficient Algorithm Level Error Detection for Number-Theoretic Transform Assessed on FPGAs.** Kasra Ahmadi, Saeed Aghapour, Mehran Mozaffari Kermani, Reza Azarderakhsh. *ACM Transactions on Embedded Computing Systems, 2025.* arXiv GitHub

7. **MAED: Mathematical Activation Error Detection for Mitigating Physical Fault Attacks in DNN Inference.** Kasra Ahmadi. *In submission.* GitHub

8. **A P2P File Sharing Market Based on Blockchain and IPFS with Dispute Resolution Mechanism.** Kasra Ahmadi, Molud Esmaili, Siavash Khorsandi. *IEEE Int'l Conf. on AI, Blockchain, and IoT (AIBThings), 2023.* IEEE GitHub

9. **[Preprint] Error Detection Schemes for τ NAF Conversion within Koblitz Curves Benchmarked on Various ARM Processors.** Kasra Ahmadi, Saeed Aghapour, Mehran Mozaffari Kermani, Reza Azarderakhsh. *TechRxiv*. TechRxiv GitHub

10. **[Preprint] Efficient Fault Detection Architectures for Modular Exponentiation Targeting Cryptographic Applications Benchmarked on FPGAs.** Saeed Aghapour, Kasra Ahmadi, Mehran Mozaffari Kermani, Reza Azarderakhsh. *arXiv:2402.18033.* arXiv

11. **[Preprint] Envisioning the Future of Cyber Security in Post-Quantum Era: A Survey on PQ Standardization, Applications, Challenges and Opportunities.** Saleh Darzi, Kasra Ahmadi, Saeed Aghapour, Attila Altay Yavuz, Mehran Mozaffari Kermani. *arXiv:2310.12037, 2023.* arXiv

TECHNICAL SKILLS

Languages: Python, C++, C, Linux/Bash

ML and LLMs: PyTorch, TensorFlow, Hugging Face Transformers, scikit-learn, LangChain, LoRA, DP-SGD, Flower, vLLM, sparse autoencoders, RAG

Security and Privacy: PII detection, privacy-preserving ML, federated learning, differential privacy, adversarial ML evaluation, fault-injection analysis, algorithm-level error detection

ML Systems: ONNX, Docker, AWS (Lambda, Glue, S3, Solutions Architect), model routing, evaluation pipelines

CERTIFICATIONS

AWS Certified Solutions Architect - Associate

LangChain for LLM Application Development

Deep Neural Networks with PyTorch (IBM)

ETL and Data Pipelines with Shell, Airflow and Kafka (IBM)

Attention Mechanisms and Transformer Models

Generative AI - Technical AI Advisor (Nvidia)

Fundamentals of LLMs (HuggingFace)

Fine-tuning Language Models (HuggingFace)

AWARDS & PROFESSIONAL SERVICE

- Best Paper Award, IEEE Security and Privacy Workshops, HMI-SA Workshop, 2025.
- Peer reviewer for 21 manuscripts across security, cryptography, embedded systems, and ML-related venues (IEEE TVLSI, TCAS-I, ACM TECS).
- Teaching assistant for Deep Learning in Computer Vision, Cryptography, Operating Systems, Computer Architecture, Network Lab, and System Design at USF.
- Mentor, REU Site: Cryptography and Coding Theory, USF (NSF Award #2244488).