# NASTARAN DARABI

Engineering Research Facility (ERF) 2024, 842 West Taylor Street, Chicago, IL
✉ ndarab2@uic.edu in linkedin.com/in/nastaran-darabi github.com/nstrndrbi G Google Scholar ☎ +1(773) 551-1451

## EDUCATION

**University of Illinois Chicago** **Jan. 2021 - March 2026**
*Doctor of Philosophy, Electrical and Computer Engineering* *Chicago, IL, USA*
*Thesis: "Ensuring Robustness and Reliability in Multimodal Autonomous Systems:*
*A Unified Framework for Multi-Level Anomaly Detection, Adaptive Resilience, and Causal Diagnostics"*

**University of Illinois Chicago** **Jan. 2024 - May 2025**
*Master of Science, Computer Science (Concurrent)* *Chicago, IL, USA*
*Natural Language Processing and Machine Learning. GPA: 4/4*

**University of Illinois Chicago** **Jan.2021 - May 2025**
*Master of Science, Electrical and Computer Engineering* *Chicago, IL, USA*
*Deep Learning. GPA: 4/4*

## TECHNICAL SKILLS

- **Languages:** Python, Matlab, Perl, SQL, NoSQL, R, C/C++, HTML, JavaScript
- **Machine Learning Libraries:** PyTorch, TensorFlow, Scikit-Learn, Pyro, Pandas
- **Technologies:** Git, Figma, Docker
- **LLM Tools and Frameworks:** Hugging Face Transformers, LangGraph, LlamaIndex, Agents, CrewAI, fedLLM
- **Relevant Courses:** (1) Machine Learning (Intro and Adv.), (2) Computer Vision (Intro and Adv.), (3) Causal Inference, (4) Natural Language Processing (Intro and Adv.), (5) Computer Algorithms, (6) Neural Networks (Intro and Adv.), (7) Reinforcement Learning, (8) Large Language Models in Production, (9) Algorithmic Fairness, (10) Database Systems

## EXPERIENCE

**Apple** **May 2025-September 2025**
***Machine Learning Intern*** *Cupertino, CA, United States*

- Design AI/ML models, and conduct model training and performance validation.
- Integrate the models and algorithms into manufacturing equipment.

**Radical AI** **April 2024-December 2024**
***Artificial Intelligence Engineer Intern*** *Remote, United States*

- Utilized cutting-edge technologies, including OpenAI and Google Gemini, to develop ReX, an AI Coach designed to be a reliable career companion for learners. ReX offers personalized coaching, mentorship, and support across different stages of their career life-cycle.

**Semiconductor Research Corporation(SRC)** **January 2023-Present**
***Research Scholar*** *Remote, United States*
Efficient multi-modality sensor fusion for 3D object detection: Conducting research on adaptive LiDAR for 3D object detection.

**University of Illinois Chicago** **January 2021-Present**
***Graduate Research Assistant*** *Chicago, IL, United States*
EigenShield: Causal Subspace Filtering via Random Matrix Theory for Adversarially Robust Vision-Language Models

- Developed EigenShield, an inference-time defense leveraging Random Matrix Theory to enhance adversarial robustness in vision-language models. In evaluations, EigenShield reduced the attack success rate by 76.5% for LLaVA-v1.5-7B and by 45.3% for MiniGPT-4, while also significantly lowering toxicity scores.

## RESEARCH INTERESTS

Vision-Language Models and Multimodal Deep Learning, Ambient LLM Agents, Adversarial Machine Learning and Robust AI, Reinforcement Learning and Intelligent Agent Design, Large Language Models, Causal Inference

## SELECTED PUBLICATIONS

- **Nastaran Darabi**, Devashri Naik, Sina Tayebati, Dinithi Jayasuriya, Amit Ranjan Trivedi, "Resilience in Ambient Multi-Agent LLMs via Decentralized Bio-Autonomic Control and Immune-Inspired Anomaly Detection", **Association for the Advancement of Artificial Intelligence Conference (AAAI, 2026), Under Review**
- **Nastaran Darabi**, Devashri Naik, Sina Tayebati, Dinithi Jayasuriya, Ranganath Krishnan, Amit Ranjan Trivedi, "EigenShield: Causal Subspace Filtering via Random Matrix Theory for Adversarially Robust Vision-Language Models", **Association for the Advancement of Artificial Intelligence Conference (AAAI, 2026), Under Review**
- **Nastaran Darabi**, Divake Kumar, Sina Tayebati, Amit Ranjan Trivedi, "INTACT: Inducing Noise Tolerance through Adversarial Curriculum Training for LiDAR-based Safety-Critical Perception and Autonomy", **IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP, 2026), Under Review**
- Davide Ettori, **Nastaran Darabi**, Sureshkumar Senthilkumar, Amit Ranjan Trivedi, "RMT-KD: Random Matrix Theoretic Causal Knowledge Distillation", **Design, Automation and Test in Europe Conference (DATE, 2026), Under Review**
- Davide Ettori, **Nastaran Darabi**, Sina Tayebati, Mahesh Subedar, Ranganath Krishnan, and Amit Ranjan Trivedi, "EigenTracks: Spectral Activation Feature Tracking for Robust and Real-Time Hallucination and OOD Detection in LLMs and VLMs", **IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP, 2026), Under Review**
- Sina Tayebati, Divake Kumar, **Nastaran Darabi**, Dinithi Jayasuriya, Ranganath Krishnan, Amit Ranjan Trivedi, "Learning Conformal Abstention Policies for Adaptive Risk Management in Large Language and Vision-Language Models", **Asian Conference on Machine Learning (ACML, 2025), Accepted**
- A.R. Trivedi, S. Tayebati, H. Kumawat, **N. Darabi**, D. Kumar, A.K. Kosta, Y. Venkatesha, D. Jayasuriya, N. Jayasinghe, P. Panda, S. Mukhopadhyay, K. Roy, "Intelligent Sensing-to-Action for Robust Autonomy at the Edge: Opportunities and Challenges", **Design, Automation and Test in Europe Conference (DATE, 2025)**
- **Nastaran Darabi**, Dinithi Jayasuriya, Devashri Naik, Theja Tulabandhula, and Amit Ranjan Trivedi, "Enhancing 3D Robotic Vision Robustness by Minimizing Adversarial Mutual Information through a Curriculum Training Approach", **IEEE International Conference on Robotics and Automation (ICRA, 2025)**
- **Nastaran Darabi**, Sina Tayebati, Sureshkumar S., Sathya Ravi, Theja Tulabandhula, and Amit R. Trivedi, "STARNet: Sensor Trustworthiness and Anomaly Recognition via Approximated Likelihood Regret for Robust Edge Autonomy", **IEEE International Joint Conference on Neural Networks (IJCNN, 2024)**
- **Nastaran Darabi**, Priyesh Shukla, Dinithi Jayasuriya, Divake Kumar, Alex C Stutts, Amit Ranjan Trivedi, "Navigating the Unknown: Uncertainty-Aware Compute-in-Memory Autonomy of Edge Robotics", **Design, Automation and Test in Europe Conference (DATE, 2024)**

## SELECTED COURSE PROJECTS

- LLM Agent for Quantum Compute Resource Optimization — CS 532 Spring 2025
- Enhancing LLM Safety with Random Matrix Theory: A Strategy for Preventing Jailbreaks — CS 521 Spring 2025
- Enhancing AI Visibility, Fairness, and Accountability — CS 516 Spring 2025
- UIC Course Planner and Catalog: build the front-end of a website Fall 2024
- Design a Specialized Chatbot from Scratch — CS 421 Fall 2024
- Ethereum Blockchain Transaction Analysis — CS 418 Fall 2024
- Bridging Causal Inference and Sensor Trustworthiness for Anomaly Recognition — CS 520 Fall 2023

## CERTIFICATIONS AND AWARDS

- Introduction to Back-End Development — Credential URL Summer 2025
- Building Transformer-Based Natural Language Processing Applications — Credential URL Fall 2023
- Supervised Machine Learning: Regression and Classification — Credential URL Summer 2022
- Advanced Learning Algorithms — Credential URL Summer 2022