

Soroush Vahidi

sv96@njit.edu | [Website](#) | [LinkedIn](#) | [GitHub](#) | [Hugging Face](#)

PROFESSIONAL SUMMARY

Computer Science Ph.D. candidate at NJIT (expected August 2027) specializing in AI/ML systems, algorithms, and optimization. Experience building production C++/Qt software, LLM serving and evaluation systems, and reproducible large-scale experimentation in Python, with additional experience in Linux, HPC/SLURM, automated testing, and validation pipelines.

EDUCATION

New Jersey Institute of Technology

Ph.D. in Computer Science

Newark, NJ

Jan. 2022 – Expected Aug. 2027

Isfahan University of Technology

B.Sc. in Computer Engineering (Computer Architecture)

Isfahan, Iran

Jun. 2015 – Jun. 2020

EXPERIENCE

Software Developer Intern — Canfield Scientific, Parsippany, NJ

Jun. 2025 – Aug. 2025

- Developed and maintained production C++/Qt medical-imaging software, building core features for multi-camera real-time image acquisition, annotation, lesion tracking, and clinical reporting.
- Debugged high-performance real-time acquisition and image-processing pipelines under strict production and memory constraints.
- Collaborated with cross-functional engineering and QA teams using Git, automated testing suites, and robust release workflows.

Graduate Research Assistant — New Jersey Institute of Technology, Newark, NJ

Jan. 2022 – Mar. 2023

- Researched, designed, and implemented high-performance parallel and combinatorial optimization algorithms for scheduling and graph problems.
- Developed and benchmarked Chapel-based parallel algorithms on HPC clusters, contributing directly to a publication in the IEEE HPEC conference.

Graduate Teaching Assistant — New Jersey Institute of Technology, Newark, NJ

2023 – Present

Supported Courses: CS 610 Data Structures and Algorithms (multiple semesters, 2023–2026), DS 677 Deep Learning (Fall 2025), CS 435 Advanced Data Structures (Fall 2025), and CS 388 Mobile Application Development / Android (Spring 2026).

- Lead labs, review student codebase and GitHub submissions, grade programming assignments, and build custom Python-based course delivery and grading automation tools.

TECHNICAL SKILLS

Languages: Python, C/C++, SQL; Java, MATLAB, Chapel

AI / ML: PyTorch, Hugging Face, vLLM, TensorFlow, scikit-learn

Data / Scientific: NumPy, pandas, SciPy

Optimization: Gurobi, CPLEX, HiGHS, MiniZinc

Systems / HPC: Linux, SLURM, CUDA, OpenMP, Chapel

Engineering: Git/GitHub, pytest, GitHub Actions/CI, FastAPI, Qt, mypy, ruff

AI Development: ChatGPT, Claude, Gemini, Perplexity, Cursor, JetBrains Junie; AI-assisted coding, debugging, testing, repository analysis, and agentic workflows

SELECTED ENGINEERING PROJECTS

LLM Serving / Scheduler Evaluation System

[GitHub](#)

Python, vLLM, SLURM

- Built reproducible LLM-serving evaluation infrastructure spanning 120 workload windows, 13 scheduler policies, and 6 calibrated operating regions (9,360 configurations, 18,720 simulator executions).
- Validated simulated findings with a 240-run real-system vLLM/GPU benchmark campaign on SLURM, utilizing custom Python telemetry instrumentation and automated validation pipelines.

NJIT Thesis Auditor

Python, FastAPI, pytest, GitHub Actions

- Built a local-first web and CLI auditing application parsing PDF, TeX, and DOCX documents with over 80 heuristic compliance checks, structured JSON/Markdown reports, SSRF safeguards, and privacy protections.
- Maintained 3,192 passing unit tests using pytest with multi-version GitHub Actions CI, clean mypy type safety, and ruff static analysis across 113 Python source files.

Retrieval-Assisted Optimization Modeling

[GitHub](#)

Python, SciPy/HiGHS

- Engineered an end-to-end Python system transforming natural-language optimization queries into structured LP/ILP formulations via retrieval-conditioned modeling.
- Validated mathematical formulations with SciPy/HiGHS solvers on a 331-query evaluation corpus, incorporating automated reproduction scripts and structured data manifests.

Learning-Augmented Cache Eviction

[GitHub](#) | [Hugging Face](#)

Python, ML, SLURM

- Developed trace-driven caching simulation and ML-prediction framework in Python, evaluating learned eviction predictors against standard baselines on HPC systems.
- Orchestrated large-scale SLURM execution campaigns, compiling and publishing the LAFC-Evict dataset (v0.3, 22.4M rows of counterfactual supervision) on Hugging Face as a reproducible public artifact.

SELECTED PUBLICATIONS

- **S. Vahidi**. “Training-Free Ranking from Pairwise Comparisons via Acyclic Graph Construction.” *The Journal of Supercomputing*, accepted/in production, 2026. [Sole Author]
- M. M. Islam, **S. Vahidi**, B. Schieber, S. B. Roy. “Promoting Fairness and Priority in Selecting k -Winners Using IRV.” *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)*, 2024. [\[DOI\]](#)
- **S. Vahidi**, B. Schieber, Z. Du, D. A. Bader. “Parallel Longest Common SubSequence Analysis in Chapel.” *Proceedings of the IEEE High Performance Extreme Computing Conference (HPEC)*, 2023. [\[DOI\]](#)

HONORS & RESEARCH SUPPORT

- **Best Paper Award**, WALCOM 2023, for “Interweaving Real-Time Jobs with Energy Harvesting to Maximize Throughput”.
- **ACM-ICPC Asia Tehran Regional** — 17th Place (2017); Honorable Mention (2015–2016); Regional Competitor (2015–2017).
- **Research Support & Grants**: Cohere Labs Catalyst Grant (2026); CloudRift AI Builder Grant (2026); Google Cloud Research Credits Program (2026).