

Zahra Taheri, Ph.D.

Applied Machine Learning Engineer | Computer Vision | ML Systems | LLMs & RAG

Tehran, Iran · zahra.taheri.ds@gmail.com

[GitHub](#) · [LinkedIn](#)

Profile summary

Applied Machine Learning Engineer with postdoctoral experience in Machine Learning & AI and 6+ years of experience across computer vision, deep learning, ML systems, and applied AI. Experienced in developing Python-based ML pipelines and integrating AI capabilities into production-facing applications, with a focus on model validation, evaluation, quality control, and reliability.

Hands-on experience with pose estimation, segmentation, human-body and musculoskeletal analysis, PyTorch, OpenCV, RTMPose, DensePose, SAM2, and MediaPipe. Strong at translating research, product, and domain requirements into modular, maintainable AI solutions while balancing model quality, robustness, processing speed, and implementation complexity.

Highly self-directed learner with a demonstrated ability to enter new technical fields independently, including transitions from mathematics into advanced graph research and later into machine learning, deep learning, and applied AI. Comfortable working independently across model development and ML engineering while collaborating with product, engineering, and domain experts.

Professional experience

AI Developer · Early-stage AI startup (remote, Germany)

February 2025 – January 2026

- Served as the sole AI engineer for a production-facing computer-vision application for musculoskeletal assessment, owning the AI workflow from problem formulation and experimentation through model development, validation, and product integration.
- Designed and developed an end-to-end computer-vision pipeline for multi-view human-body and musculoskeletal assessment, processing user-captured mobile images into structured anatomical measurements using pose estimation, segmentation, landmark refinement, and geometry-based post-processing.
- Developed modular Python pipelines integrating RTMPose, DensePose, SAM2, MediaPipe, OpenCV, and related open-source models to improve landmark localization, segmentation robustness, and measurement consistency across body views.
- Built validation and quality-control components for human detection, view classification, confidence estimation, and anatomical plausibility checks to improve reliability under imperfect input conditions.
- Worked directly with physicians, orthopedists, and sports-medicine experts to translate clinical and product requirements into AI features, measurements, validation rules, anomaly-analysis logic, and outputs supporting personalized corrective-exercise recommendations.
- Evaluated practical trade-offs among model accuracy, inference reliability, processing speed, and pipeline complexity when selecting and integrating computer-vision approaches.
- Worked with the FastAPI-based backend and Swagger/OpenAPI interfaces to test and validate AI API workflows and support integration of model outputs into the production application, collaborating with the software engineer responsible for backend and server deployment.

AI Systems Engineer · LLM Systems Lab (independent open-source project)

January 2026 – September 2026 · [Repository](#)

- Designed and implemented a modular Retrieval-Augmented Generation system covering document ingestion, embedding-based indexing, semantic retrieval, context construction, and configurable local/remote model inference.
- Developed benchmark-driven evaluation and regression-testing infrastructure using human-verified datasets and retrieval metrics (Hit Rate@K, Recall@K, MRR) to diagnose failures and support continuous improvement.
- Implemented structured logging, latency tracking, metadata capture, unit testing, and reproducible configuration to improve debugging, observability, maintainability, and system reliability.

Senior Research Fellow · Machine Learning and Graph Mining Lab, Shahid Beheshti University

January 2020 – February 2025 · [MLGM Lab](#)

- Independently developed, trained, benchmarked, and evaluated machine-learning and deep-learning models for graph-based representation learning and molecular property prediction.
- Developed FunQG, a quotient-graph-based framework that reduced Graph Neural Network computational complexity through graph coarsening while preserving predictive performance.
- Conducted knowledge-graph representation-learning experiments using PyKEEN, including entity-relation modeling, knowledge-graph embeddings, and link prediction.
- served as a supervisor and advisor to multiple M.Sc. students on knowledge graphs, GNNs, and LLM-based AI systems.

Postdoctoral Research · Machine Learning & AI, Shahid Beheshti University

May 2021 – April 2022

- Conducted postdoctoral research on Graph Neural Networks and deep learning for healthcare applications, focusing on representation learning, predictive modeling, and explainable AI.
- Transitioned from mathematics into ML and deep learning through intensive self-study and hands-on experimentation, building the foundation for subsequent applied AI work.

Research Assistant · Institute for Research in Fundamental Sciences (IPM)

September 2017 – August 2018

- Conducted research in computational algebraic graph theory, graph invariants, and combinatorics, contributing to mathematical modeling and peer-reviewed publications.

Education & research training

Postdoctoral Research · Machine Learning & Graph AI, Shahid Beheshti University

May 2021 – April 2022

Focused on graph neural networks and deep learning for healthcare applications, representation learning, predictive modeling, and explainable AI.

Ph.D. in Mathematics · University of Isfahan

September 2014 – December 2019

Graduated with distinction · GPA 19.25/20 (approximately 3.85/4.0) · Ph.D. thesis grade: 20/20.

Academic & teaching experience

- Invited Lecturer in graduate-level Graph Neural Networks & Machine Learning.
- Teaching Assistant in Applied Machine Learning, Shahid Beheshti University.
- Co-Instructor in graduate-level Optimization for Data Science, Shahid Beheshti University.
- served as a supervisor and advisor to multiple M.Sc. students on knowledge graphs, GNNs, and LLM-based AI systems.

Selected publication

FunQG: Molecular Representation Learning via Quotient Graphs · Shared first author.

Hajiabolhassan, H.; Taheri, Z.; Hojatzia, A.; Yeganeh, Y. T. *Journal of Chemical Information and Modeling* 2023, 63(11), 3275–3287. <https://doi.org/10.1021/acs.jcim.3c00445>

Technical skills

- **Programming & ML engineering:** Python; PyTorch; scikit-learn; Git/GitHub workflows (branches, pull requests, collaborative development); unit testing; regression testing; configuration-driven development; structured logging.
- **Computer vision & deep learning:** RTMPose, DensePose, SAM2, MediaPipe, OpenCV; pose estimation; segmentation; landmark detection/refinement; human detection; multi-view human-body analysis; confidence estimation; geometry-based measurements; quality control.
- **ML systems & evaluation:** Modular ML/AI pipelines; production-facing AI integration; model benchmarking and validation; evaluation frameworks; latency measurement; metadata tracking; failure analysis; reproducible experimentation; practical performance/reliability trade-offs.
- **Machine learning:** Deep learning; representation learning; Graph Neural Networks; predictive modeling; hyperparameter optimization (Optuna, Ray Tune); RNN/LSTM fundamentals; hands-on educational time-series forecasting project using COVID-19 case-count data.
- **Development environment:** Linux/Ubuntu (4 years of regular use); Bash; software/package installation; SSH remote-server workflows; Windows 11 with WSL2; basic SQL concepts and common commands.
- **Additional AI:** RAG systems; embeddings; semantic retrieval; LLM inference; knowledge graphs; link prediction; PyKEEN; Transformer fundamentals; LangChain; ChromaDB; Ollama; Gemini.
- **Technical exposure — sports video analytics:** Independent study of football video-analytics concepts including player detection, multi-object tracking, ReID, tracklet-level identity resolution, trajectory refinement, team assignment, jersey OCR, identity fragmentation, and ID-switch evaluation.

Current skill development

Continuing to build practical capability in FastAPI, Docker, workflow orchestration, n8n automation, Agentic AI, and MCP. FastAPI experience documented above is limited to validating AI API workflows through Swagger/OpenAPI in collaboration with the backend engineer; these areas are not presented as advanced specialties.

Professional skills

Independent technical ownership, rapid self-learning and self-directed problem solving; translating domain-expert and product requirements into practical AI/ML solutions; cross-functional collaboration with engineering, product, medical/domain, and research teams; AI evaluation, benchmarking, failure analysis, technical documentation, and knowledge sharing.

Languages

Persian: Native · English: Upper Intermediate / Professional Working Proficiency (CEFR B2+)