📍 Tehran, Iran
🌐 GitHub | Medium | LinkedIn | Website

Alireza Amouzad

📞 +98 938 196 1903
✉️ a.r.amouzad.m@gmail.com

Summary

Results-driven **Machine Learning Engineer** 3+ years of experience designing, developing, and deploying scalable ML systems. Skilled across the full ML lifecycle (research, development, deployment) with expertise in PyTorch, HuggingFace, LangChain, Docker, and FastAPI. Improved user engagement (+20%) and retention (+15%) through optimized AI integration. Built deep learning models for tumor segmentation and classification in mammography, enhancing cancer detection accuracy and supporting clinical decisions. Passionate collaborator dedicated to solving complex problems through innovative and maintainable machine learning .

Experiences

Sharif AI Center Tehran
Machine Learning Engineer Aug 2023 - Mar 2025

- Designed and deployed 10+ end-to-end AI microservices for a super-app. Services included **real-time TTS, ASR, OCR, NER, language detection, and image analysis (age/gender)**.
- Drove a 20% uplift in user **engagement** and 15% increase in user **retention** by designing, and integrating these novel AI features, collaborating closely with backend, frontend, and product teams.
- Engineered scalable inference pipelines using FastAPI and Docker.

AI-Med Tehran
Intern and AI Developer Dec 2021 - Jun 2023

- Developed and fine-tuned deep learning models for **tumor segmentation in mammography images**, achieving high accuracy in delineating suspicious regions.
- Built classification pipelines on segmented areas to **differentiate malignant from benign tumors**, enhancing cancer detection performance and supporting clinical decision-making..

Allameh Helli High School Tehran
Teacher of Programming Jul 2017 - Sep 2019

- Delivered engaging Python programming and algorithm courses, fostering improved **problem-solving** abilities and a passion for computer science among high school students.
- Mentored students through hands-on projects, leading to notable achievements such as winning national coding competitions, with several students going on to earn gold medals in both national and international Olympiads in Informatics.

Skills

- **Languages** Python
- **Data Analysis** NumPy, Pandas, Matplotlib, Seaborn
- **ML framework** Scikit-Learn, TensorFlow, PyTorch, LangChain
- **DevOps** Docker
- **Model Serving** FastAPI
- **Software Engineering** Git, Unit/Integration test (Pytest), Documentation

Education

Amirkabir University of Technology Tehran
Master of Science, Artificial Intelligence Oct 2017 - Sep 2021

- Thesis: Developed **GIUNet**, a novel Graph Isomorphism U-Net architecture introducing **pqPooling**, a unique graph pooling method. This innovation addressed key limitations in graph representation learning.
- Achieved **state-of-the-art results** on benchmark graph classification datasets (MUTAG, PTC, ENZYMES), demonstrating significant improvements over existing GNN models.
- Published findings in the high-impact journal **Expert Systems with Applications (2024)**.

Rajaei University Tehran
Bachelor of Science, Electrical Engineering Oct 2011 - Sep 2016

Publications

Graph Isomorphism U-Net
Expert Systems With Applications 2024

- GIUNet combines U-Net's hierarchical architecture with graph isomorphism convolution and introduces pqPooling(in spectral and centrality regime) to enhance graph classification. By integrating node features and graph structure, GIUNet achieves state-of-the-art performance on benchmark graph classification datasets, including MUTAG, PTC, and ENZYMES, demonstrating superior discriminative capabilities.

Projects

Image Similarity Search with MinIO Storage

- **Summary:** A production-ready system for multimodal image similarity search using SigLIP2 + CLIP embeddings, integrated with Milvus for high-speed vector retrieval and MinIO for scalable S3-compatible storage. Exposed via FastAPI endpoints and an interactive Gradio UI, it supports async embedding, batch indexing, and presigned image access. Fully containerized for distributed environments.
- **Key Technologies:** Python, FastAPI, Gradio, PyTorch, Transformers, Milvus, MinIO, Docker.

Persian ghazal criticizer agent

- **Summary:** An agent for automated critique of Persian ghazals using parallel LangChain pipelines, accessible via FastAPI endpoints and Gradio UI; managed through Poetry and dotenv, leveraging Pydantic for structured outputs and containerized with Docker for streamlined, scalable deployment.
- **Key Technologies:** Python, FastAPI, Gradio, LangChain, Pydantic, Docker.

GIUNet

- **Summary:** Implemented a Graph Isomorphism Network for graph classification using PyTorch, featuring novel models and a custom pooling method (pqPool). The project includes scripts for data loading, training, and evaluation, supporting multiple datasets with detailed training metric logging.
- **Key Technologies:** Python, PyTorch, Networkx, PyG

Connect Four and Pentago with Reinforcement Learning

- **Summary:** Developed a Board games Connect Four and Pentago where a reinforcement learning agent learns optimal strategies through self-play. Utilized deep Q-learning techniques to enhance the AI's gameplay, integrating a graphical user interface with PyQt for an engaging user experience.
- **Key Technologies:** Python, PyTorch, Gym, PyQt