AmirMasoud Azadfar

Montreal, Canada • +1 343 988 3995 • amirmazadfar@gmail.com • LinkedIn • GitHub • ORCID • Portfolio

PROFESSIONAL SUMMARY

AI Systems Engineer with 5+ years of experience designing and deploying production-grade intelligent systems, combining large language models, agentic workflows, and scalable microservice architectures. Proven track record of building end-to-end AI platforms, including real-time APIs, distributed pipelines, and data-intensive backend systems. Strong expertise in Python-based service design, REST APIs, message-driven architectures, and cloud-native deployment, with hands-on ownership of reliability, scalability, and production operations.

TECHNICAL SKILLS

- **Languages:** Python, SQL, C/C++, Bash, R
- **AI / Machine Learning:** PyTorch, TensorFlow, Scikit-learn, Stable-Baselines3, HuggingFace Transformers, Sentence-Transformers, spaCy, MLflow, Weights & Biases
- **LLM / Retrieval / Search:** RAG, CAG, Agentic Workflows, LangChain/LangGraph, Qdrant, Elasticsearch, BM25, Semantic Search, Vector Embeddings
- **Data Extraction & Automation:** Scrapy, Playwright, Selenium, BeautifulSoup, OCR, Pandas, NumPy
- **Backend & Systems:** FastAPI, Redis, Kafka, Docker, Nginx, REST APIs, Microservices
- **Data & Infrastructure:** PostgreSQL, MySQL/MariaDB, MongoDB, Neo4j, CI/CD, Git, OpenTelemetry, Prometheus, Sentry, GCP, AWS

WORK EXPERIENCE

Machine Learning Developer March 2023 – Present
CanApply (NovaVidya Inc.) – Montreal, Canada (Full-time)

- Architected and deployed a production-grade, microservice-based AI platform composed of a provider-agnostic LLM client, DAG-based orchestration runtime, and FastAPI services, enabling scalable multi-agent workflows via RESTful APIs and service-oriented architecture (41 domain-specific operators).
- Designed a contract-driven operator framework using JSON manifests, Jinja2 prompt templates, and JSON Schema validation, enabling hybrid deterministic/LLM execution, policy enforcement, human-in-the-loop gates, and resilient operator-level retries, circuit breakers, and rollout controls.
- Built and deployed a distributed conversational AI system with 20+ integrated tools, exposing real-time inference via streaming APIs (SSE) and supporting multi-provider LLM execution across OpenAI, Anthropic, and Gemini.
- Engineered a production-scale backend system for automated outreach, integrating recommendation pipelines, multi-agent workflows, and asynchronous task processing with external APIs (Gmail OAuth) and persistent state management.
- Built an AI-powered faculty crawling and digestion platform (CanSpider - Professors) that replaced manual collection workflows, expanding coverage from 18,000 professors across 350 departments and 31 institutions in 4 months to 64,127 professors across 2,325 departments and 99 institutions in 40 days through LLM-generated crawl plans, Scrapy/Playwright execution, Kafka pipelines, and review tooling.
- Developed internal operational interfaces using Next.js (React, Tailwind CSS) to manage AI workflows, orchestration pipelines, and system state, enabling real-time control and monitoring of production processes.
- Designed and implemented event-driven pipelines using Kafka for asynchronous data ingestion, processing, and coordination between AI services, improving system scalability and fault tolerance.
- Developed a second AI extraction platform for academic program intelligence (CanSpider - Programs), combining autonomous URL discovery, static/dynamic rendering detection, LLM-based structured extraction, OCR fallback, confidence-weighted aggregation, and config-driven enrichment pipelines for tuition, deadlines, admissions, and language requirements.
- Engineered an AI-powered academic program recommendation system combining semantic vector search with dynamic rule-based multi-filter matching via Qdrant. Designed FastAPI endpoints for real-time student-to-program personalization.
- Built and deployed AI data products including an immigration news intelligence system (CanNews) integrating 23 sources, OpenAI-powered summarization, Elasticsearch semantic search, REST APIs, and Telegram delivery pipelines.
- Owned end-to-end system deployment and operations, including Dockerized microservices, Nginx routing, CI/CD pipelines, cloud provisioning (AWS/GCP), monitoring and alerting, caching layers (Redis), and production reliability.

Software Engineer — Quant & Machine Learning Dec 2017 – Aug 2023
Sepanta Communications Technology Co.

- Built large-scale financial data pipelines aggregating technical, fundamental, and news data for 400+ Tehran Stock Exchange equities; extended infrastructure to crypto markets (Binance, KuCoin) with real-time WebSocket ingestion (¡500 ms latency).
- Designed algorithmic trading strategies (mean reversion, momentum, order-book microstructure) and implemented derivatives pricing models including Black–Scholes and binomial trees.
- Developed machine learning models for market segmentation using consensus clustering (K-Means, hierarchical) on historical financial time series.
- Built a Persian (Farsi) NLP sentiment analysis pipeline by labeling 15K+ financial text samples and training TF-IDF + Naive Bayes classifiers used in production analytics.
- Architected a Telegram-based financial analytics platform serving 10K+ users with REST APIs, asynchronous processing, real-time trading signals, and sentiment-driven alerts.
- Developed automated trade execution modules, realistic backtesting environments, and a high-frequency crypto triangular arbitrage system scanning 1000+ trading pairs.
- Deployed and maintained production systems using FastAPI, MongoDB, Redis, Docker, and AWS EC2.

SELECTED PROJECTS

Trellion – AI Hiring Intelligence Platform (Founder / ML Engineer) – trellion.ai 2025–2026
Building an AI-native B2B SaaS hiring platform that replaces manual resume screening and early-stage interviews with automated evaluation pipelines. Developing LLM-based interview analysis, candidate ranking systems, and workflow automation tools that transform job descriptions and candidate responses into structured hiring signals and role-fit scoring.

AI-Powered Pet-Caregiver Matching System – pets.pawsome.vip 2025
Deployed a production-grade FastAPI backend to match pet-care service requests with qualified caregivers. Combined multi-layer rule-based filtering with a structured 22-question rubric evaluated by an LLM agent for compatibility scoring. Integrated Redis, MySQL, OpenAI + Google APIs, and Sentry.io. Powering 100+ monthly bookings on the public-facing concierge platform.

Chronos-Powered Crypto Forecasting & Trading System – GitHub 2025–2026
Built an end-to-end BTC/USDT futures forecasting and trading pipeline combining Amazon Chronos-2, LightGBM quantile models, regime-aware strategy logic, realistic cost modeling, and walk-forward backtesting/paper-trading for deployment evaluation.

RL-Driven Blackjack Simulator with Action-Masking – GitHub 2025
Designed and benchmarked a Gym-compatible blackjack environment with partial observability and legality-aware actions. Trained PPO and DQN agents using Stable-Baselines3; DQN achieved a 49.4% loss rate, outperforming the house edge.

GNN-Based Pharmacological Interaction Engine – Research Project – GitHub 2024
Built a graph neural model to predict drug interaction severity using SMILES-based molecular graphs and RDKit descriptors. Modeled 16,000+ pharmacological structures with GAT, GIN, and MPNN variants to benchmark predictive performance.

RESEARCH & PUBLICATIONS

Amini, S. *et al.*, **Azadfar, A.** – "Comprehensive Compilation and Quality Assessment of Street-Level Urban Air Temperature Measurements Across European Networks." – *Scientific Data (Nature Portfolio, Springer Nature)*, 2026. – **DOI**

Khatibi, S., Rahmani, A., **Azadfar, A.**, Fonseca, V. P. da, and Oliveira, T. E. A. – "ViTHL: Vision Transformer-Based Hybrid Localization for Humanoid Robots." – *RoboCup Symposium 2025*, Salvador, Brazil. – **vithl.github.io**

EDUCATION

Honours Bachelor of Science in Computer Science
Lakehead University, Thunder Bay, Canada

HONORS & AWARDS

Ahwazi Young Investigator Award – Behavioral Neuroscience 2016
3rd Place – National Cognitive Neuroscience Competition (Hosted by IPM) 2016
4th Place – Sharif University Robotics Competition (Smart Gardeners League) 2012