# Mohammad-Reza Azizi

+44 7383 377216 | mrazizi.work@gmail.com | LinkedIn | GitHub

## Experience

**MyTrek AI** — Masqat, Oman (Remote)
*Senior AI Engineer* — *Mar 2025 – Feb 2026*

- Contributed as one of two core developers building MyTrek, an **AI-powered learning assistant platform** for personalized education through conversational AI, automation, and browser-based workflows.
- Built backend and AI infrastructure using **Python**, **FastAPI**, **LangChain**, **LiteLLM**, **OpenAI**, **Gemini**, **OpenRouter**, **PostgreSQL**, and **Redis**, including authentication, payments, AI usage tracking, cost calculation, and core platform APIs.
- Designed AI orchestration services including **RAG pipelines**, chatbot memory, context retention, system prompts, and optimized conversational prompting strategies to improve personalization and response quality.
- Collaborated directly with the CEO on product direction, feature planning, prompt design, and technical prioritization within a lean engineering team.

**Ceres Technology** — Austin, USA (Remote)
*AI Engineer* — *Mar 2024 – Mar 2025*

- Built and trained **LightGBM** models on supply-chain data to estimate **delay risk**, and turned the results into clear charts that helped business teams understand patterns and make better decisions.
- Developed a supply-chain analytics **chatbot** using **LangChain**, **Vanna.ai**, **OpenAI**, and **Plotly**, allowing users to upload files, query external sources, combine datasets, generate visualizations, and retrieve relevant news.
- Designed backend services with **FastAPI**, **PostgreSQL**, **AWS Fargate**, and **AWS Step Functions** to support data processing, model workflows, and automated jobs in a scalable way.
- Built a **dataset recommendation service** that compared user data against more than **300k datasets** and returned relevant results in under **15 minutes** using efficient multiprocessing.

**Core Decentralized Technologies** — Bratislava, Slovakia (Remote)
*AI Engineer* — *Mar 2023 – Mar 2024*

- Built a content-based video recommendation system using **LLM-generated embeddings**, storing vectors in **Qdrant** and finding related videos with **KNN**. Automated the pipeline with **Apache Airflow** and exposed it through **FastAPI** with **Redis** caching.
- Developed an **awareness detection** module for a KYC workflow using **MediaPipe** head-pose estimation, improving the accuracy of user verification.
- Expanded into broader software engineering by creating **Python** microservices with **PostgreSQL** and **Redis**, implementing service-to-service communication with **gRPC** and **Apache Kafka**, managing database migrations with **Alembic**, and writing thorough **unit and integration tests**.

**Digikala**
*Data Scientist* — *Feb 2021 – Mar 2023*

- Optimized a trend-finding algorithm by adding **parallel processing** across the product catalog, achieving a **1200% speedup** in offline processing tasks.
- Built a complementary recommendation system using **KNN** on image and text embeddings stored in **Qdrant**, increasing **add-to-cart rate by 70%** and **daily GMV by 202%**, validated through a diff-in-diff experiment.
- Developed popular-product and shop-recommendation services by combining **trending** and **top-selling** signals, and exposed the results through **FastAPI** APIs with **Redis** caching.
- Processed large-scale datasets with **Apache Spark** and scheduled recurring offline workflows using **Apache Airflow**.

**Part AI Research Center**
*Data Scientist / Speech Chapter Lead* — *Jan 2020 – Feb 2021*

- Improved an **Awareness Detection using Speech** microservice by analyzing model errors and adding new features to make the system more reliable.
- Researched and experimented with state-of-the-art **ASR** models in English, testing different architectures and training approaches.
- Created and maintained **text and speech datasets**, including data collection, cleaning, denoising, and quality control.
- Trained and evaluated speech models using **TensorFlow**, **PyTorch**, **ESPNet**, **Nvidia NeMo**, and **Optuna**, while leading a team of four data scientists through task planning, code reviews, and daily coordination.

## Education

**Amirkabir University of Technology**
*M.Sc. in Artificial Intelligence* — *2019 – 2021*

**Ferdowsi University**
*B.Sc. in Computer Engineering* — *2014 – 2019*

## Skills

**AI/LLMs:** LangChain, LiteLLM, CrewAI, RAG, multi-agent systems, OpenAI API, Gemini API, Vanna.ai, mem0, Tavily

**ML/Data:** LightGBM, PyTorch, TensorFlow, Scikit-learn, PySpark, Spark, embeddings, recommendation systems

**Backend:** Python, FastAPI, PostgreSQL, Redis, gRPC, REST

**Databases:** PostgreSQL, MySQL, MongoDB, Qdrant, Redis, SQLAlchemy, Alembic

**Cloud/DevOps:** AWS ECS, Fargate, Lambda, Step Functions, S3, CloudWatch, Docker, Git, CI/CD

**Data/Workflow Tools:** Airflow, Kafka, n8n

**Testing:** Unit tests, integration tests, TestContainers