

# Shayan Ekramnia

## Machine Learning Engineer

Naples, Italy | +39 351 853 3047 | shayan.ekramnia1@gmail.com  
github.com/shayan-ekramnia | linkedin.com/in/shayan-ekramnia-8b2133299

### PROFESSIONAL SUMMARY

Machine Learning / AI Engineer with 5+ years designing, training, and deploying production-grade ML and AI systems end to end — from data engineering and ETL through model development, MLOps, serving, and monitoring. End-to-end architect and developer of a complete production AI system for national retail, built from scratch and now rolling out across the branch network. Deep expertise spanning time-series and probabilistic forecasting (including fine-tuned foundation models), deep learning, multi-agent reinforcement learning, computer vision, and NLP/speech, grounded in advanced statistical and causal analysis. Strong in Python, PyTorch, and GPU-accelerated containerized pipelines, with a track record of translating business requirements into reliable, measurable ML outcomes. M.Sc. in Data Science (**110/110 cum laude**), University of Naples Federico II.

### TECHNICAL SKILLS

**Machine Learning:** PyTorch, TensorFlow, Scikit-Learn, gradient boosting (XGBoost, LightGBM, CatBoost), statistical learning, causal inference, hyperparameter optimization (Optuna)

**Time-Series Forecasting:** ARIMA/SARIMA, exponential smoothing (ETS), Prophet, neural forecasting (TFT, N-BEATS, DeepAR), time-series foundation models (Chronos), hierarchical forecasting & reconciliation, probabilistic & quantile forecasting, conformal prediction, time-series feature engineering, backtesting (rolling-origin CV)

**LLMs & Generative AI:** Hugging Face Transformers, LLM fine-tuning (LoRA/PEFT), retrieval-augmented generation (RAG), vector databases & embeddings, semantic search

**Reinforcement Learning & Computer Vision:** PPO, DQN, A3C, multi-agent RL, Gymnasium, OpenCV, YOLO, DeepFace, CoreML

**NLP & Speech:** Whisper (ASR), SpeechBrain (ECAPA-TDNN), spaCy, full-text search, Neo4j graph NLP

**MLOps & Deployment:** Docker, Apache Airflow, MLflow, model registry, experiment tracking, drift monitoring, CI/CD (GitHub Actions), FastAPI, distributed & GPU-accelerated training and inference

**Data & Cloud:** Python (expert), SQL, Bash, Apache Spark, Apache Kafka, ETL/ELT pipelines, PostgreSQL, MySQL, Neo4j, Redis, AWS & cloud ML platforms, Linux

**Visualization & BI:** Plotly, Matplotlib, Streamlit, Tableau, Power BI

### PROFESSIONAL EXPERIENCE

#### Machine Learning Engineer / AI Engineer

Nov 2024 – Present

SIACLOUD (TEKNOTEAM S.R.L.), Naples, Italy

- Single-handedly designed, built, and deployed a complete production AI system — data engineering, ML engineering, and MLOps — for a retail network of **110+ branches**, replacing fully manual, spreadsheet- and experience-driven planning.
- Ship **20,000+ probabilistic (quantile) forecasts daily** across **7,000+ SKU-level time series**, owning the entire pipeline — ingestion and ETL of a 50M+-row (and growing) dataset, training, deployment, serving, and dashboards.
- Fine-tune time-series foundation models (Chronos-Bolt) alongside classical models, with continuous automated evaluation of per-pipeline forecast accuracy on rolling held-out test windows.
- Perform advanced exploratory and causal analyses of company-wide retail data to identify demand drivers, quantify effects, and inform both model design and commercial decision-making.
- Automated the ML lifecycle with Apache Airflow, MLflow, and Docker: reproducible unattended retraining, versioned model registry, one-click rollback, and GPU-accelerated containerized training and inference across staging and production.
- Drive the production rollout in partnership with infrastructure and network teams, with Streamlit/Plotly monitoring of model performance, data drift, and forecast KPIs in near real time.

#### Machine Learning Engineer & Software Developer

Jun 2018 – Oct 2022

Vira Vision Co., Esfahan, Iran

Selected through a university elite-talent program for placement at a high-tech park company; worked alongside B.Eng. studies until 2020, then full-time.

- Engineered an AI-based COVID-19 risk-prediction system deployed across **18 companies and 28,000 employees** during the pandemic.
- Built and deployed ML models for customer-behaviour prediction across the company's SaaS products, informing retention campaigns and product analytics.
- Developed Python SaaS modules (CRM, ERP, supply-chain) with embedded analytics and ML-driven recommendations for enterprise clients.
- Built automated data pipelines, feature-engineering workflows, and BI dashboards (Tableau, Power BI) for marketing, sales, HR, and finance analytics, with post-deployment monitoring and A/B testing.

SELECTED PROJECTS

Multi-Agent Reinforcement Learning for Warehouse LGV Optimization2025 – 2026

- Built a complete multi-agent RL system to optimize routing and task sequencing for automated guided vehicles (LGVs), with a custom Gymnasium-compatible environment featuring kinematic constraints, collision avoidance, and dynamic pallet positioning.
- Implemented PPO, DQN, and A3C from scratch in PyTorch and benchmarked them against four classical planners (A\*, Dijkstra, Greedy, Random) over 100 evaluation episodes.
- PPO ranked first overall — efficiency score 78.4/100 vs. 65.3 for A\*** — with an 87.5% task-completion rate and 1.2 collisions per episode; gains validated with Welch's t-tests ( $p < 0.001$ ) and large effect sizes (Cohen's  $d > 0.8$ ).
- Delivered a Streamlit dashboard with real-time simulation, statistical significance testing, and learning-curve analysis.

Real-Time Facial Emotion Detection System2025

- Engineered a real-time emotion-recognition pipeline using DeepFace and OpenCV, supporting both 7-class (Ekman) and 3-class valence outputs.
- Implemented an internal HTTP API exposing 5-second rolling aggregates, and benchmarked five face-detector backends (RetinaFace, MTCNN, MediaPipe, SSD, OpenCV) for accuracy-latency trade-offs.

Voice-Enabled Quote Retrieval Chatbot (NLP + Speech)2025

- Designed and built an end-to-end voice NLP system over a 1.3M-quote Wikiquote corpus stored in Neo4j with full-text indexing.
- Integrated Whisper ASR, SpeechBrain ECAPA-TDNN speaker identification (192-dim embeddings), and TTS for a personalized multi-user voice experience.
- Built fuzzy author resolution and top-K ranked retrieval with normalized similarity scoring and intent routing.

On-Device iOS ML Apps — Banalize & EyeLightApple Developer Academy, 2025

- Banalize: built an iOS app using CoreML to assess fruit ripeness and classify bread types with privacy-preserving on-device inference.
- EyeLight: built an assistive-vision app using CoreML and the Vision framework for real-time object detection, distance estimation, and obstacle alerts for visually impaired users.

Neighborhood Analytics ModelFater S.p.A. & UNINA, 2024

- Developed an accrual model to estimate the market potential of new stores in the baby-diaper category, informing the company's market-expansion strategy.

EDUCATION

M.Sc. Data Science — Machine Learning & Data Analysis2023 – Oct 2025

University of Naples Federico II (UNINA), Naples, Italy

- Final grade: **110/110 cum laude**, completed alongside full-time industry work. Industry thesis: the end-to-end AI demand-forecasting system designed and deployed in production at SIACLOUD.
- Coursework: Advanced Machine Learning, Data Mining, Big Data Technologies, Statistical Learning, Ethics of AI.

Apple Developer Academy Diploma — iOS, UX/UI & AI IntegrationSep 2024 – Jun 2025

University of Naples Federico II, Naples, Italy

- Developed AI-integrated iOS apps using CoreML and the Vision framework; pitched concepts in cross-functional startup teams.

B.Eng. Mining EngineeringSep 2015 – Mar 2020

Isfahan University of Technology (IUT), Esfahan, Iran

- Head of the Engineering Scientific Association (2018–2019). Coursework in Computational Methods, Statistics, and Linear Algebra.

AWARDS & CERTIFICATIONS

- National Elite — Iran's National Elites Foundation (INEF), 2020: national selection recognizing top-performing graduates and young talents.
- Professional Project Management — QAL UK Academy, Dec 2022.
- Business Administration — University of Tehran, Sep 2023.

LANGUAGES

English (Fluent) | Persian (Native) | Italian (Intermediate)