

Ali Asfahani

Beirut, Lebanon | ali.asfahani.lb@outlook.com | +961 76 415 001 | LinkedIn | GitHub

PROFESSIONAL SUMMARY

AI Engineering master's student with hands on experience building AI applications, RAG systems, LLM agents, computer vision pipelines, and FastAPI services. Passionate about transforming AI concepts into practical products by combining machine learning, backend engineering, and user focused design. Through academic projects and bootcamps, I developed production style systems spanning retrieval, evaluation, orchestration, and deployment.

SKILLS

Programming: Python, Java, SQL, OOP, Design Patterns, REST APIs, Agile Development
AI / Data: Computer Vision, NLP, Fine-Tuning, RAG, Model Evaluation, Drift Detection, Scikit-learn, PyTorch, OpenCV
LLMs / Agents: Prompt Engineering, Embeddings, Semantic Search, LLM Prompt Chaining, Tool-Calling Agents
Backend / Infra: FastAPI, Redis Queue, PostgreSQL, pgvector, SQLAlchemy, Docker, HashiCorp Vault
Web: HTML, CSS, JavaScript, PHP, React

EXPERIENCE

Freelance Software Developer Feb 2023 – Dec 2024

- Delivered custom software and web solutions for clients, translating business requirements into maintainable applications.
- Built and modified applications using Java and web technologies, improving reliability, usability, and client workflows.
- Worked directly with clients to diagnose issues, propose solutions, and iterate based on feedback.

Web Development Intern Nov 2022 – Jan 2023

- Developed a dynamic web platform for an education center using HTML, CSS, JavaScript, PHP, and asynchronous communication.
- Designed core system workflows, database models, and user roles to support platform functionality.
- Implemented role-based access control and contributed within an agile development process.

PROJECTS

Maintainer's Copilot *NLP, Advanced RAG, LLM Agents, React*

- Built an AI assistant that classifies software issues, extracts entities, summarizes discussions, and answers questions using documentation and resolved issues.
- Implemented hybrid retrieval, reranking, metadata filtering, query rewriting, and a tool-calling chatbot with evaluation checks.

Document Classifier Service *ConvNeXt, FastAPI, SFTP Pipeline, HashiCorp Vault*

- Fine-tuned ConvNeXt models on RVL-CDIP document images to classify 16 document types with production-style evaluation.
- Built a secure FastAPI service with JWT authentication, RBAC, Redis caching, Alembic migrations, SFTP ingestion, MinIO storage, and RQ-based inference.

Drift Triage Co-Pilot *MLOps, LangGraph, FastAPI, Redis*

- Built an MLOps assistant that detects data drift, generates triage reports, and recommends actions using agent-based workflows.
- Integrated MLflow, FastAPI, Redis queues, Postgres checkpointing, and human approval gates for production-sensitive changes.

Concierge AI SaaS *Multi-Tenant AI, RAG, FastAPI, PostgreSQL*

- Developed a multi-tenant AI concierge platform with CMS-managed knowledge, an embeddable website assistant, lead capture, and human escalation.
- Built tenant-safe retrieval and access control using PostgreSQL RLS, tenant-filtered vector search, signed widget tokens, and security tests.

Spectra AI – Screen-Aware Cross-Platform Copilot (In Development) *RAG, LLM Agents, FastAPI, Computer Vision*

- Developing a screen-aware AI copilot that combines screenshots, browser DOM data, uploaded files, and official documentation to provide contextual guidance and analytics.
- Building the MVP with RAG, multimodal analysis, authenticated chat, usage tracking, and automated Excel report generation.

EDUCATION

MS – Artificial Intelligence 2025 – 2027
Université Saint Joseph (USJ) – ESIB

BS – Computer Science 2020 – 2023
CNAM University

BOOTCAMPS

SE Factory – AI Engineering Bootcamp
Forward MENA – Full-Stack Development with AI

LANGUAGES

Arabic (Native), English (Fluent), French (B2)