

# MED AMIN GNICHI

AI Engineer | Generative AI | RAG | LLM Systems | Applied Machine Learning

Remote

gnichimedamine@gmail.com | linkedin.com/in/med-amine-gnichi | github.com/medamine1920

## PROFESSIONAL SUMMARY

AI Engineer with hands-on experience building production RAG systems, document-processing pipelines, and LLM-powered decision tools across legal, financial, enterprise, and supply chain domains. Comfortable owning end-to-end work, from ingestion and retrieval to evaluation, caching, and deployment, across both self-managed Python stacks and Azure-hosted enterprise environments.

Currently freelancing on a production legal RAG platform; background spans enterprise AI tooling on Azure and inventory forecasting in supply chain operations. Looking for AI / ML engineering roles where production reliability matters as much as model quality.

## TECHNICAL SKILLS

- **Programming:** Python, SQL
- **AI & LLM Systems:** RAG, LLMs (OpenAI / Azure OpenAI / Google Gemini), Embeddings, Prompt Engineering, LangChain, LLMOps, Model Evaluation
- **Vector Search & Retrieval:** FAISS, AstraDB, Vector Databases, semantic caching
- **ML & Data:** scikit-learn, regression, Pandas, NumPy
- **Backend & APIs:** FastAPI, REST APIs, Docker, Authentication / RBAC
- **Infrastructure & Tools:** Azure, Azure OpenAI, Cassandra, Git, GitHub, Monitoring / Logging
- **Document Processing:** OCR pipelines, multi-format ingestion (PDF, image, Excel)

## PROFESSIONAL EXPERIENCE

### .FLOW

Mar 2025 - Present

#### Freelance AI Engineer

- Reduced legal-document lookup time by ~55% by deploying a production RAG chatbot using FastAPI, e5-large-v2 embeddings, Gemini, and semantic retrieval.
- Expanded retrieval to scanned legal documents by building an OCR-based PDF ingestion pipeline for files standard parsers could not process.
- Reduced duplicate LLM calls by implementing semantic caching with FAISS and Cassandra for similar user queries.
- Protected sensitive legal content by implementing multi-user sessions, RBAC, admin audit visibility, Docker deployment, and application logging.
- Improved retrieval and prompting quality by building an evaluation framework using F1 scoring across 500+ labeled questions with user-feedback-driven labels.

### Impactive

April 2023 - June 2024

#### Founding Data Lead

- Reduced wasted media spend by 18% across \$5M+ annual ad spend by building regression models measuring media-efficiency decay.
- Predicted 90-day returning revenue with 95% accuracy by developing customer cohort models.
- Reduced peak-period stockouts by 40% and improved stock-level prediction accuracy by 90% by building forecasting models across three brands.
- Cut decision-making time by 45% by owning reporting and dashboard infrastructure and iterating with Growth Strategists.
- Reduced onboarding time by 60% and eliminated 250+ hours of annual manual work by building an internal AI knowledge system and automated reporting workflows.

### Alight MEA

Oct 2022 - Mar 2023

#### AI Engineer - Intern

- Supported an enterprise AI platform across Germany, MENA, and Tunisia by building on Microsoft Azure and Azure OpenAI Service.
- Enabled retrieval across PDF, image, and Excel files by building a multi-format document-ingestion pipeline.
- Grounded chatbot responses in uploaded documents by implementing chunking, embeddings, and semantic retrieval for the RAG pipeline.
- Integrated RAG components into backend workflows supporting real-time interaction and preparation for a multi-office production rollout.

### SagemCom

Feb 2021 - June 2021

#### ML Engineer - Intern

- Reduced inventory stockouts by 25% by deploying linear and ridge regression models to predict reorder points.
- Improved model reliability by building automated monitoring pipelines and engineering features across multi-source supply-chain data.

### Cipi-Actia

Jan 2020 - Feb 2020

#### Software Developer – Intern

- Built a quality monitoring platform ("Quality Wall") to detect and classify production defects across the factory line, enabling searchable defect records and statistical reporting.
- Designed the database architecture and classification logic to categorize defect types, providing production teams with structured insights to identify recurring issues and improve line efficiency.

## EDUCATION

### Bachelor of Science in Data Science Engineering With Honors

Graduated Nov 2025

Private College of Engineering & Technology - TEK-UP

### Associate Degree in Information Systems Development With Honors

Sep 2018 - July 2021

Higher Institute of Technological Studies - ISET