

ALEXANDRE FERNANDES

AI/ML Software Engineer

Schleswig, Germany (commutable to Hamburg) | EU Citizen | +49 177 4896716 | alexandre@alexandrefernandes.dev
alexandrefernandes.dev | github.com/AlexPT2k22 | linkedin.com/in/alexandrerofernandes

PROFILE

Computer Science graduate specializing in AI/ML engineering, with hands-on experience building production-grade RAG pipelines, agentic systems, and LLM evaluation frameworks. Delivered a Text-to-SQL system that improved retrieval accuracy from 13% to 77%, using a custom LLM-as-Judge framework to detect hallucinations and validate output quality. Comfortable working without established tutorials or off-the-shelf solutions, building from first principles when the problem has no existing blueprint.

EXPERIENCE

Geodouro

02/2026 – 06/2026

AI Engineer Intern

- Led the development of a production RAG + Text-to-SQL system using Qwen2.5 7B, improving accuracy from 13% to 77% on the development set (72% on the generalization set).
- Designed an agentic retrieval pipeline with LangGraph and BM25+MMR hybrid search, covering 500+ tables across 3 PostgreSQL databases.
- Built a domain-specific LLM-as-Judge evaluation framework to detect hallucinations and measure output quality, validating retrieval strategies across a 50-query golden set.
- Implemented dynamic few-shot prompt engineering and a self-correction/validation loop, outperforming a 14B baseline at 5.2s median latency, with end-to-end observability via Langfuse.

EDUCATION

Escola Superior de Tecnologia e Gestão de Viseu (ESTGV - IPV)

10/2023 – 06/2026

Bachelor of Science (BSc), Computer Sciences Engineering

- Final Grade: 15/20 (Equivalent to German 2.5).
- Final Curricular Internship & Project: 20/20 (Equivalent to German 1.0).

PROJECTS

Zyllo

01/2026 – Present

AI Solutions Engineer

- Designing a SaaS platform that automates TikTok direct messages and follow-ups for e-commerce stores using an LLM-driven conversational engine.
- Defined end-to-end architecture (database schema, auth, LLM orchestration) and TikTok API integrations for automated, context-aware replies.

AI Smart Parking System

11/2025 – 01/2026

AI Engineer (Computer Vision)

- Built an intelligent parking system using a custom-trained CNN (PyTorch/ResNet) for real-time spot occupancy detection and Fast-ALPR for automatic license plate recognition with ESP32 IoT cameras.
- Developed a FastAPI backend with PostgreSQL, WebSocket real-time updates, Supabase image storage, and a React Native mobile app for end-user access.

Enterprise Training Platform (Softinsa / IBM Subsidiary)

02/2025 – 06/2025

Academic Partnership Project – Full-Stack Developer

- Built a full enterprise training platform (Web/Mobile/Backoffice) with JWT authentication, course management, scalable architecture using React, Node.js/Express, and PostgreSQL.

SKILLS, LANGUAGES & CERTIFICATIONS

AI, ML & Computer Vision: LLMs, RAG, Agentic Systems (LangGraph), Prompt Engineering, LLM-as-Judge Evaluation, Fine-tuning (Unsloth/QLoRA), LangChain, Langfuse, SQL Agents, Ollama, ChromaDB, Whisper.

Backend & Data Engineering: Python, Node.js, JavaScript, SQL, FastAPI, REST APIs, n8n.

Databases & Infrastructure: PostgreSQL, Supabase, MongoDB, Docker, Git.

Frontend & UI/UX: React, Flutter, Figma.

Languages: Portuguese (Native), English (Fluent - C1), German (Beginner - A1, actively learning).

Certifications: German Language (BTU Cottbus-Senftenberg), Intro to Git and GitHub (Google).