

BEATRIZ MAIA

Machine Learning Engineer | Computer Vision & GenAI

Guarapari, ES, Brazil

✉ contact@beamaia.com

☎ +55 27 996480747

🐙 github.com/beamaia

🌐 [linkedin.com/in/beamaia](https://www.linkedin.com/in/beamaia)

🆔 orcid.org/0000-0003-1057-2785

Professional Summary

Machine Learning Engineer with 5+ years of ML experience across research and industry, currently building production computer-vision systems for livestock monitoring, with a background in GenAI, LLM fine-tuning, and cloud ML workflows. With C2 fluency in English and Portuguese, A1 French, and multicultural experience from living in both Brazil and the United States, I adapt well to diverse, global teams.

Skills and Expertise

- **Computer Vision:** object detection, instance and semantic segmentation, stereo vision, depth estimation, 3D point clouds (PointNet/PointNet), YOLO, OCR.
- **Generative AI and Agents:** Google ADK, MCP/FastMCP, Agno, CrewAI, LangGraph, LangChain, agent/tool orchestration, Claude Code and OpenAI Codex, prompt engineering and context management, RAG/CAG, embeddings, LLM-as-a-judge evaluation.
- **LLM Training and Optimization:** DeepSpeed, FSDP, Accelerate, SFT, RLHF/DPO, AWQ quantization, multi-GPU scaling.
- **ML Frameworks:** PyTorch, Transformers, TensorFlow, Keras, scikit-learn, NER.
- **Cloud and Production:** GCP, Vertex AI, Cloud Run, Cloud Build, Azure, Docker, NVIDIA Triton, model serving, Git, GitLab CI/CD.
- **Backend and Data:** FastAPI, PostgreSQL, SQLAlchemy, Pandas, PyTest.
- **Languages:** English (C2), Portuguese (C2), French (A1).

Experience

Olho do Dono

May/2026 – Present

Machine Learning Engineer

Brazil

- Develop computer-vision models for cattle monitoring from video, covering object detection, instance segmentation, and multi-object tracking for automated herd counting.
- Build cattle weight-prediction regression models from stereo (ZED) and monocular images, using depth estimation and 3D point-cloud learning.
- Build CV training workflows and run data analysis: detector experiments and dataset auditing (tiling, annotation conversion, bounding-box validation).
- Deploy models into production applications using Docker and NVIDIA Triton Inference Server.
- Create AI assistant skills and RAG pipelines to automate internal workflows and streamline documentation retrieval.

Tech Stack: Python, PyTorch, FastInst, YOLO, SAM 2, LocateAnything, OpenCV, NVIDIA Triton, Docker, GCP, Vertex AI

Cadastra (contracted to Grupo Boticário)

Apr/2025 – Mar/2026

Machine Learning Engineer

Brazil

- Maintained a GenAI Gateway via Azure/Vertex AI and extended an SDK with LiteLLM for unified model routing across Agno, CrewAI, and Google ADK workflows.
- Developed FastAPI services for RAG-as-a-service and transcription, integrating advanced prompt engineering and MCP/FastMCP for tool-enabled workflows.
- Automated a GCP data pipeline (Cloud Run/Build), reducing update times from one hour to five minutes, and implemented LLM-as-a-judge evaluation pipelines.

Tech Stack: Python, FastAPI, LiteLLM, Google ADK, Agno, CrewAI, MCP/FastMCP, LangChain, Vertex AI, Cloud Run, Cloud Build, PostgreSQL

Weni

Nov/2023 – Oct/2024

Machine Learning Engineer

Brazil

- Fine-tuned models up to 70B parameters using PyTorch, DeepSpeed, FSDP, and SFT/RLHF/DPO for zero-shot classification and Q&A tasks.
- Applied AWQ quantization to deploy a 70B model in place of a 7B model, improving task performance by approximately 15% without increasing inference latency.
- Compared and evaluated RAG pipelines utilizing LangChain, hybrid search, Cohere reranking, and Amazon Bedrock.

- Developed semi-autonomous RunPod.io training pipelines, increasing weekly experiment capacity from 2 to 15.

Tech Stack: Python, PyTorch, DeepSpeed, FSDP, Accelerate, SFT/RLHF/DPO, AWQ, LangChain, Amazon Bedrock

Bio-inspired Computing and Engineering Lab (LABCIN)

Aug/2020 – Oct/2024

Researcher

Brazil

- Lead author of research on Transformers and Few-Shot Learning applied to digital pathology and clinical skin images, published in *Expert Systems with Applications* (2023).
- Performed rigorous data curation on high-variance clinical skin imaging datasets, identifying and mitigating data leakage and contamination risks.
- Investigated Variational Autoencoders (VAEs) for medical feature extraction; peer-reviewed for the IEEE Journal of Biomedical and Health Informatics.

Tech Stack: Python, PyTorch, CNNs, Transformers, SIFT, VAEs, medical imaging, data curation

Eldorado Research Institute

Jun/2022 – Oct/2023

Machine Learning Engineer

Brazil

- Developed a hybrid OCR pipeline using OpenCV and Tesseract, with YOLO for document-element detection.
- Developed NLP and Named Entity Recognition pipelines using Hugging Face for unstructured text extraction.
- Curated Python-based Docker images to reduce vulnerability risk in ML environments; participated in MLOps and Explainable AI research groups.

Tech Stack: Python, PyTorch, OpenCV, Tesseract, YOLO, Hugging Face, NLP/NER, Docker, MLOps

VERT Transformação em TI

Jun/2021 – Jun/2022

Data Scientist (Intern to Full-time)

Brazil

- Led client meetings and developed risk-analysis frameworks for government bidding and DETRAN-ES financial projects.
- Architected ETL processes in SAS (Base SAS, SAS Viya) and managed database mapping with PostgreSQL and SQLAlchemy; developed BeautifulSoup and Selenium scrapers for public data collection.

Tech Stack: Python, SAS 9.4, SAS Viya, PostgreSQL, SQLAlchemy, BeautifulSoup, Selenium

Education

Federal University of Espírito Santo (UFES)

Expected Dec/2026

Bachelor's Degree in Computer Science

Vitória, ES, Brazil

Coursework completed in 2024; currently finishing thesis research. Emphasis on Machine Learning and Software Engineering.

Volunteering

Introcomp – Introduction to Computing

Aug/2019 – Nov/2022

Teacher

Vitória, ES, Brazil

Taught Python to public high school students, created Jupyter-based study materials, led data collection and class-streaming committees, and co-authored a paper on programming education challenges presented at the 2021 Brazilian Symposium on Computing Education.