Mario Parreño

+34605170713 | maparla@proton.me | Valencia, Spain

Summary

Machine Learning engineer with a proven track record in Computer Vision, NLP, and Generative AI. Passionate about leveraging Deep Learning to solve real-world problems and deliver innovative solutions that save companies time and money. Skilled in transforming complex challenges into practical applications, I am dedicated to developing cutting-edge AI systems that showcase my expertise in automation and problem-solving. My goal is to continue advancing in the field, creating impactful solutions that push the boundaries of artificial intelligence.

Experience

Newfire Global Partners
Senior Machine Learning Engineer | 10/2024
Lead Machine Learning Engineer for a strategic internal project designed to automate talent discovery and resource allocation through agentic AI.

- **Architected a multi-agent orchestration layer** using **LangGraph** and **LangChain**, implementing state-aware workflows that autonomously decompose complex project requirements into searchable technical criteria.
- **Engineered a hybrid retrieval system** that combines **LLM-driven keyword extraction** with **semantic vector search**, significantly increasing expert recall by mapping natural language intent to specific technical skill sets.
- **Developed an automated ranking and sorting engine** that evaluates candidate-to-task alignment, providing explainable relevance scores and prioritized lists based on granular CV analysis and historical project data.
- **Integrated comprehensive tracing and observability** into the agentic lifecycle, leveraging **LangSmith** to monitor LLM outcomes, quantify retrieval precision, and rigorously test for "hallucinations" in candidate profiles.

Senior Machine Learning Engineer at Sway AI, an innovative American startup developing a no-code AI platform. My main duties have been:

- Design, optimize, and deploy advanced prompts to drive effective interactions with **LLMs across diverse AI use cases**, enhancing platform performance and user experience.
- Built a Retrieval-Augmented Generation (RAG) **pipeline for a Due Diligence automation solution**, enabling accurate and scalable responses to over 300 questions using LLMs.
- Conducted **research on chunking techniques and embedding models for RAG workflows**, **reducing latency by 10×** while maintaining high retrieval and generation accuracy.
- Evaluated and compared OCR technologies, recommending improvements that significantly enhanced the quality of LLM-based document processing. Explored **alternative OCR techniques** that improved overall LLM pipeline accuracy by over 8%.
- Implemented a **parallel processing pipeline** using multithreading, increasing overall application throughput by more than 2.2×.

Cognizant
Senior Machine Learning Engineer | 09/2022 - 10/2024

- **Transformer Fine-Tuning & Extraction:** Fine-tuned **Transformer** architectures (**Donut** encoder-decoder) using **Hugging Face** and **PyTorch** for end-to-end entity extraction from complex legal and financial documents; leveraged **Transfer Learning** to adapt models to domain-specific vocabularies and low-quality scans.
- **MLOps & Data Evaluation Leadership:** Orchestrated production pipelines on **Azure ML** leveraging **BigQuery** for complex event data storage; led a team of 20 using **Label Studio** and **Vertex AI** (500k+ annotations) to perform rigorous error analysis and **inter-annotator agreement** checks, effectively mitigating **concept drift**.
- **Document Classification:** Engineered multi-stage classification systems using **scikit-learn** to automate document triage and metadata extraction, improving downstream pipeline efficiency.
- **Generative AI PoCs:** Developed and evaluated tool-based conversational agents using **LLMs**, focusing on prompt engineering and the creation of custom evaluation frameworks to measure model reasoning and performance.
- **Constrained CV Solutions:** Designed computer vision models optimized for consumer-grade hardware to detect mobile app components, implementing strict **data privacy protocols** to prevent leakage.

Solver Intelligent Analytics
Deep Learning Engineer | 04/2022 - 09/2023

- Led an interurban road maintenance project by implementing **AI-driven predictive analytics**.

- Streamlined the entire pipeline from business **problem analysis, data labelling, model experimentation, reporting and production**. Reduced time to deployment and accelerated project delivery.
- Developed **multi-label classification models** and damage segmentation algorithms to enable proactive maintenance planning.
- Pioneered stereoscopic depth-mapping technology to enhance **3D defect categorisation** and improve repair planning efficiency.
- Achieved road defect detection and segmentation accuracy superior to human performance, resulting in a overall cost reduction for the client.

Polytechnic University of Valencia
Machine Learning Engineer | 04/2018 - 01/2022

- Deep learning researcher, focusing on computer vision, image translation, segmentation, and binary and multi-class classification.
- Developed an AI-based asphalt damage assessment system for the construction sector, reducing manual inspection time and improving accuracy. Enabled data-driven maintenance decisions to increase road maintenance efficiency by prioritising critical repairs and optimising resource allocation.
- Pioneered the application of CycleGAN models to medical imaging, achieving state-of-the-art accuracy in cross-domain adaptability of diagnostic algorithms at an international workshop.

Education

Polytechnic University of Valencia | Valencia, Spain
Master's degree | 09/2018
Artificial Intelligence, Pattern Recognition and Digital Imaging

Polytechnic University of Valencia | Valencia, Spain
Bachelor's degree | 09/2017
Informatics Engineering, Computer Science

Skills

Deep Learning, Machine Learning, Data Science, Team Management, LLMs, Generative AI, Computer Vision

Languages

Spanish, English, Catalan; Valencian

Publications

DASeGAN: Domain Adaptation and Generalization for Medical Segmentation Tasks via Generative Adversarial Networks

Proposed a method that maps images into a universal domain, unifying the image appearance across different sources. The approach is based on the CycleGAN architecture, using an image generator trained to produce task-specific, realistic images with indistinguishable domains of origin.

Deidentifying MRI Data Domain by Iterative Backpropagation

Investigated novel domain adaptation procedure for medical image analysis to address challenges in multi-vendor, multi-center data distribution variations. Developed a method using classifier training and gradient-based image modification to adapt models to unseen data distributions.

A deep analysis on high-resolution dermoscopic image classification

Comprehensive analysis of the effectiveness of state-of-the-art deep learning techniques when applied to dermoscopic images. To achieve this goal, several CNN architectures are analysed, measuring how their performance is affected by network size, image resolution, data augmentation process, amount of available data, and model calibration.

Links

- Aidventure: My personal blog about the latest trends in Deep Learning. Coded and evaluated some of the most popular architectures and techniques of LLMs for image and text processing. Fine-tuned LLMs models for sequence and token

classification, question answering, instruction following, and more.

- **GitHub**: For my side projects and my blog, I have more than 50 repositories where I share my experiences with Computer Vision, LLMs, RAG and diffusion models.
- **LinkedIn**: A more complete view of my background, including my certifications on various platforms, honors and awards.
- **Kaggle**: Competed in classification, recognition and identification problems. Expert member for 7 years. Ranked in the top 3% competitions and 8% as dataset contributor.