Guillermo Caminero Fernández

ML Engineer | Production Systems & Pipeline Optimization

Madrid, Spain   gcaminerof@gmail.com   +34 653 579 976   GitHub: gcamfer

LANGUAGES

Spanish (Native)

English (Professional)

KEY SKILLS

Computer Vision
- Image Classification
- Object Detection
- Object Segmentation

ML & MLOps
- Classification & Regression
- Clustering
- Reinforcement Learning
- Time Series Forecasting

Data Engineering
- ETL Pipelines
- PySpark
- Databricks

Geospatial
- Geospatial Data Science
- ADAS Features

PROGRAMMING
- Python (8+ yrs)
- Java (1 yrs)
- C/C++ (1 yrs)
- Go

PROFESSIONAL SUMMARY

M.Sc. in Telecommunications Engineering with extensive experience in machine learning, computer vision, and NLP. Started working mainly in natural language processing but developed a strong passion for computer vision, leading to expertise in facial recognition, object detection, and image segmentation. Currently focused on geospatial data science and maps analytics for ADAS features at TomTom. Passionate about putting ML models into production and optimizing ETL pipelines to achieve the best throughput while minimizing costs. Over the past 2 years, I have specialized in profiling and optimizing Python code to reach the maximum potential of each workflow, ensuring efficient and scalable data processing systems. Great passion for electronics, embedded systems (RaspberryPi, Arduino, FPGAs), cybersecurity, and pentesting. Always looking for new challenges to reach full potential.

Total Experience: 7+ years

PROFESSIONAL EXPERIENCE

Senior Data Scientist
TomTom • Madrid, Spain   Jul 2022 - Present

Maps Analytics: • Developing metrics to analyse the quality of maps for ADAS attributes enabling map releases to be 100% compliant with quality standards • Stakeholder orientation to focus data-driven efforts • Extraction of ADAS features from sensors using ML enabling quality checks without manual intervention saving 60% of the pipeline cost and 90% of the time delay for the automatic metrics • Designed a unified pipeline for metric creation that reduced the time to create a new metric from 6 months to 4 weeks • Geospatial data science from indexing, databases, operations and analysis • Software architectural designs for scaling the metrics pipeline and fallback mechanisms MLOPS: • Productivization of ML models in Databricks (CV detection and classification) scaling to 10k+ jobs per day • Dockerizing the ETL and inference environments for production ready deployments • ETL Load Optimization for lidar data and images into indexed delta lake reducing the time to load the data 60% • ETL design for download and process aerial lidar from goverment agencies • Monitoring the performance and costs from baseline to production alignign with VM reservations and cost targets • Performance optimization of the pipeline reducing the baseline cost by 80% using rust and cython based libraries and optimizing the code AI Agents: • Automatic feedback classification using LLMs and RAG for reducing manual effort of low value feedback by 40% • [POC] AI agent for querying the map (talking to the map), using claude + rag + smolagents

Technologies: PySpark, Python, Databricks, Geospatial Data Science, ADAS, MLOps

Senior Data Scientist
Telefónica • Madrid, Spain   Aug 2020 - Jul 2022

Main projects: • Sport analytics for first division football team, video summary generation, player similarity detection, spotfire dashboards • Data science professor (internal training) for data science and machine learning from basics data wrangling to pyspark machine learning models • Segmentation and classification of aerial drone imagery for defect detection • Time series forecasting • EDA and KPI visualizations in Spotfire for non-profit foundation • Auto EDA and auto ML tool for internal use • Research on facial recognition system for video door entry systems, proof of life and detecting onboarding frauds

Technologies: Python, PySpark, TensorFlow, Plotly, Dash, OpenCV, Spotfire, Docker, MySQL, FastAPI, FFmpeg

NLP/NLU Analyst
Minsait (Indra) • Madrid, Spain   Aug 2018 - Aug 2020

Document management department: • Metadata extraction from banking and legal documents, email information • Text classification & sentiment analysis • Computer vision: Object detection (signatures, stamps, brands, logos) • OCR image preprocessing • REST services development (Flask, Spring Boot)

Technologies: Java, Python, Flask, TensorFlow, Keras, NLTK, Spacy, Regular Expressions, Tesseract, Abbyy FineReader

Fundamentals of MCP
Hugging Face
Jul 2025

AI Agents Fundamentals
Hugging Face
Jun 2025

Databricks Certified Data Engineer Professional
Databricks
Mar 2025

Databricks Certified Machine Learning Associate
Databricks
Dec 2024

Microsoft Certified: Azure Fundamentals
Microsoft
Mar 2024

Microsoft Certified: Azure Data Scientist Associate
Microsoft
Jul 2021

Extracurricular Internships in Research Group
Apr 2016 - May 2016

GIR de Electrónica (Universidad de Valladolid) • Valladolid, Spain

• Electronic design of a system capable of performing the electrical-optical characterization of laser diodes • Manufacture and assembly of the system developed on printed circuit board (PCB)

Technologies: *PCB Design,   Electronic Design,   Laser Characterization*

Internships
Feb 2016 - Apr 2016

GIR de Electrónica (Universidad de Valladolid) • Valladolid, Spain

• Reverse engineering of the Spectra Pro 300i monochromator

Technologies: *Reverse Engineering,   Optical Systems*

EDUCATION

Master in Telecommunication Engineering
Oct 2016 - Jul 2018

Universidad de Valladolid • Valladolid, Spain

Grade: 8.213/10

- Signal processing/analysis (Matlab and Python)
- Programming FPGAs (Verilog)
- Web development (HTML, CSS, Javascript, PHP, MySQL)

Degree in Engineering Specific Telecommunication Technologies
Sep 2013 - Jul 2016

Specialization: Electronic Systems

Universidad de Valladolid • Valladolid, Spain

Grade: 7.0/10

- PCB design (Proteus Design Suite)
- Electronic analysis & Electronic communications
- Electromagnetic compatibility
- Microcontroller programming (ARM7, PIC)

PUBLICATIONS

Adversarial environment reinforcement learning algorithm for intrusion detection. *Computer Networks* (2019).