Pablo Catret

Data Scientist with strong quantitative skills and hands-on experience applying **AI** to real-world challenges in sectors such as climate and agriculture. Results-driven and an effective communicator, focused on transforming **data** into actionable **business value.** Keen to support innovation and decision-making through tech.

CONTACT INFORMATION

E-Mail: pablo.catret@gmail.com
LinkedIn: pablo-catret
Personal Website

SKILLS

Data & AI:
Python, R, Machine Learning, Deep Learning (Tensorflow, PyTorch), Statistics, Optimization, pySpark

Cloud & Automation:
AWS, Airflow, Docker, FastAPI, Git

Databases & Visualization:
SQL, Neo4J, MongoDB, Cassandra, PowerBI

Business & Communication:
A/B Testing, Causal Inference, Financial Analytics (KPIs), Project Communication, Agile teamwork and technical presentations to stakeholders

Generative AI & NLP:
LLMs, LangChain, RAG, Autonomous Agents, Webscraping, HuggingFace Transformers

LANGUAGES

Spanish & Valencian: Bilingual (C1 JQCV)
English: Fluent (C2 EF SET, B2 PACLE)
German: Intermediate (B1 EOI)
Italian: Intermediate

EXPERIENCE

Data Science Researcher
IRTIC
Sept. 2024 – Present
- Responsible for the complete data cycle (ETL, modelling, visualisation) in several research projects.
- Development of ML/DL models for classification, prediction and anomaly detection in different domains.
- Automation of data pipelines and design of solutions with modern data science tools.
- Co-author of a paper published in Q1 journal (IJAEOG) and several papers under review.

Data Science Intern
Capgemini
Feb. 2024 – Jun. 2024
- Contributed to an Anti-Money Laundering (AML) fraud detection project using Explainable ML and Graph ML.
- Worked on an applied ML project in smart cities and infrastructure monitoring.
- Hands-on training in Large Language Models (LLMs) and Generative AI with a focus on enterprise integration.
- Strengthened communication skills through cross-functional collaboration.

Data Science Research Collaborator
IRTIC
Jan. 2023 – Sept. 2024
- Built and evaluated ML/DL models for prediction and time-series forecasting in environmental science.
- Co-authored scientific articles and participated in international research initiatives.

EDUCATION

MSc in Data Science - University of Valencia

Sept. 2024 – Present · GPA: 9.48/10

BSc in Mathematics - University of Valencia

Sept. 2020 – Jul. 2024 · GPA: 8.63/10

MAIN CERTIFICATES:

Stanford University: Machine Learning Specialization
NYIF: Machine Learning for Trading Specialization
Yale University: Financial Markets (with honors)
AWS: Generative AI with Large Language Models
John Hopkins University: Advanced Statistics for Data Science Specialization
Politecnico di Milano: Project Management Specialization
IBM: AI Workflow: Feature Engineering and Bias Detection; Scalable Machine Learning on Big Data using Apache Spark

OTHER RECENT PROJECTS:

- **ML** pipelines in **AWS** for real-time inference (2025)
- **Dockerized Airflow pipeline** for financial time-series, anomaly detection, and persisting processed data in PostgreSQL (2025)
- **AI assistant agent** for farmers, integrating **LLMs** and agronomic data (Currently)
- **Author** of "**A/B Testing**: Guía Práctica para Resultados Reales" (2025)