



# Enrique Pardo García

CIENTÍFICO DE DATOS E INGENIERO DE DATOS

✉ [enrique.pardo.ab@gmail.com](mailto:enrique.pardo.ab@gmail.com) | 📞 +34 641 385 553 | 📍 A Coruña, España  
🌐 [enriquepardogarcia.com](http://enriquepardogarcia.com) | [in linkedin.com/in/enrique-pardo-garcia](https://www.linkedin.com/in/enrique-pardo-garcia) | [github.com/EnriquePardoGarcia](https://github.com/EnriquePardoGarcia)

## RESUMEN PROFESIONAL

Científico de Datos e Ingeniero de Datos en formación en la Universidade da Coruña, con sólida experiencia en modelado estadístico y aprendizaje automático. Especializado en desarrollar soluciones analíticas de extremo a extremo, combinando ingeniería de datos con investigación aplicada en machine learning. Con experiencia profesional como Ingeniero de Datos en ENXENIO S.L., donde optimicé pipelines de datos y modelos analíticos en entornos de producción. Galardonado con **2.º lugar** en el *Datathon 2026* por diseñar una plataforma integral de inteligencia marítima con seis modelos especializados de aprendizaje automático. Apasionado por investigar nuevas técnicas de machine learning, modelado estadístico riguroso y entregar soluciones que generen valor comercial tangible en entornos innovadores.

## EXPERIENCIA PROFESIONAL

*Ingeniero de Datos en Prácticas*  
ENXENIO S.L. · A Coruña, España

Septiembre – Octubre 2025

- Validación y auditoría de métricas comerciales clave mediante expresiones DAX y consultas PostgreSQL, garantizando integridad de datos en todas las capas de información.
- Documentación y mantenimiento de modelos de datos en entornos Power BI, mejorando la transferencia de conocimiento y eficiencia de incorporación.
- Optimización de paneles de control a través de visualizaciones avanzadas y configuraciones de Seguridad a Nivel de Fila (RLS) adaptadas a necesidades de interesados.
- Auditorías de datos sistemáticas basadas en SQL para detectar y resolver inconsistencias en conjuntos de datos en producción.
- Diseño e implementación de navegación de informes dinámicos utilizando marcadores en Power BI, mejorando la experiencia del usuario final.

## EDUCACIÓN

*Grado en Ciencia e Ingeniería de Datos*  
Universidade da Coruña, A Coruña, España

2022 – Presente

*Bachillerato Tecnológico* (Especialidad STEM)  
IES Alonso Berruguete, Palencia, España

2015 – 2021

## CERTIFICACIONES Y GALARDONES

- Aprendizaje Profundo con PyTorch para Análisis de Imágenes Médicas** — Udemy
- 2.º Lugar, Datathon 2026** — Instituto Tecnológico de Galicia & Universidade da Coruña

## PROYECTO DESTACADO

- Diseñé e implementé un pipeline de seis etapas procesando 171.750 incidentes marítimos sintéticos a través de 5.725 embarcaciones, con más de 170 características de ingeniería de atributos.
- Construí seis modelos especializados de aprendizaje automático: detección de anomalías (XGBoost, 97,6%), clasificación de riesgos (Random Forest, 92,6%), clasificación de tipo de incidente (XGBoost, 99,9%), regresión de costos (Gradient Boosting,  $R^2 = 0,70$ ), pronóstico temporal (Prophet) y clustering geográfico (HDBSCAN).
- Desarrollé un panel de control web completo con Next.js 14, TypeScript, Tailwind CSS, Recharts y Leaflet, presentando KPIs ejecutivos, centro de acciones, explorador de incidentes, mapas de calor interactivos y seguimiento de embarcaciones.

## HABILIDADES TÉCNICAS

**Lenguajes de Programación** | Python R TypeScript SQL DAX Bash

**Librerías de Ciencia de Datos y Machine Learning** | Scikit-learn XGBoost LightGBM CatBoost Pandas NumPy SciPy Statsmodels

**Redes Neuronales y Visión por Computadora** | PyTorch TensorFlow OpenCV Albumentations Torchvision

**Series Temporales y Análisis** | Prophet Statsmodels Pandas SciPy

**Clustering y Análisis Exploratorio** | Scikit-learn HDBSCAN Scipy

**Bases de Datos** | PostgreSQL SQL MongoDB

**Procesamiento Distribuido y Big Data** | PySpark Apache Spark Pandas

**Procesamiento de Lenguaje Natural** | Lucene Scikit-learn Pandas

**Visualización e Inteligencia Empresarial** | PowerBI Recharts Leaflet Matplotlib Seaborn Plotly

**Desarrollo Web Frontend** | Next.js 14 React TypeScript Tailwind CSS HTML5 CSS3

**Control de Versión y Gestión de Código** | Git GitHub GitHub Actions

**DevOps y Containerización** | Docker CI/CD

**Herramientas y Entornos de Desarrollo** | Jupyter Notebook JupyterLab VS Code R Studio Overleaf  $\LaTeX$

## OTROS PROYECTOS ACADÉMICOS Y PERSONALES

- **Segmentación de Fluido Patológico en OCT** — Suite experimental de ablación con 5 experimentos para segmentación binaria de fluido patológico en imágenes de Tomografía de Coherencia Óptica (OCT). Cada experimento aísla la influencia de una variable (aumento de datos, función de loss, encoder preentrenado). Implementación de K-Fold validación cruzada (5 folds), early stopping, scheduler adaptativo y post-procesado morfológico (closing, opening, drop CCs). Evaluación completa con métricas Dice, IoU, Precision, Recall, AUC-ROC, curvas ROC y PR. Reproducibilidad garantizada con seed global determinístico. Tecnologías: **PyTorch, UNet, ResNet34, Albumentations, 5-Fold CV, funciones de loss (BCE, Dice, Focal), post-procesado morfológico, visualización comparativa.**
- **Clasificación Visual de SmartPorts** — Clasificación binaria de imágenes para vigilancia portuaria marítima. Detección de embarcaciones desde imágenes aéreas e identificación de barcos en puerto mediante afinamiento de modelos. Comparación de entrenamiento desde cero frente a transferencia de aprendizaje desde ImageNet, analizando el impacto de la proporción de validación y el aumento de datos en precisión, F1 y AUC-ROC. Tecnologías: **PyTorch, ResNet18, aumento de datos, validación cruzada.**
- **Segmentación de Carreteras desde Imágenes Aéreas** — Pipeline de segmentación semántica utilizando modelos de aprendizaje profundo en imágenes de satélite de alta resolución. Tecnologías: **Redes Convolucionales, PyTorch, segmentación semántica.**
- **Reconocimiento de Objetos en Imágenes** — Sistema de clasificación de imágenes basado en redes neuronales convolucionales, entrenado y evaluado en conjuntos de datos visuales de referencia. Tecnologías: **CNN, PyTorch, visión por computadora.**
- **Sistema de Recuperación de Información Clínica** — Motor de recuperación basado en Procesamiento de Lenguaje Natural para consultas inteligentes de registros médicos estructurados y no estructurados. Tecnologías: **NLP, Lucene, BM25, embeddings, TREC.**
- **Detección de Malware en IoT con PySpark** — Pipeline de detección de anomalías distribuido para tráfico de redes IoT, construido sobre Apache Spark para procesamiento a gran escala. Tecnologías: **PySpark, Big Data, procesamiento distribuido, detección de anomalías.**

## IDIOMAS

**Español** — Lengua materna

**Inglés** — C1 Avanzado (Cambridge Certificate)

## COMPETENCIAS CLAVE

- Pensamiento Analítico y Crítico
- Colaboración Multidisciplinaria
- Proactividad e Iniciativa Independiente
- Atención Rigurosa al Detalle
- Adaptabilidad Rápida a Nuevas Tecnologías
- Comunicación Clara y Estructurada