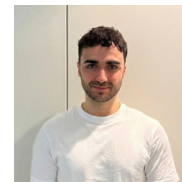


Iñaki Martín Sorribes

Estudiante de Ciencia de Datos · Co-Fundador @ Iceberg Datum Analytics · Valencia, España

imxrtins@gmail.com · [imartinsorribes.github.io](https://github.com/imartinsorribes) · linkedin.com/in/inaki-martin-sorribes · github.com/imartinsorribes



PERFIL

Estudiante de Ciencia de Datos en la Universitat de València y co-fundador de Iceberg Datum Analytics, startup de IA que ofrece soluciones de Machine Learning y NLP para los sectores legal y sanitario. Experiencia desplegando sistemas en producción sobre AWS. Busco prácticas para septiembre de 2026 como ML/NLP engineer o data scientist, abierto a fintech, retail, industrial, salud o energía.

EXPERIENCIA

Co-Fundador — Iceberg Datum Analytics S.L.

Nov 2025 — Actualidad

Valencia, España · Media jornada

- Co-responsable del desarrollo de negocio: captación de clientes, gestión del pipeline comercial y negociación de contratos bajo NDA.
- Co-lidero el desarrollo técnico de Iceberg Health (plataforma de datos sanitarios) e Iceberg Law (procesamiento documental basado en LLMs).
- Diseño y despliegue de infraestructura en producción sobre AWS usando Docker y FastAPI.
- Actualmente ejecutando pilotos bajo NDA con clientes potenciales en ambos sectores.

Investigador en prácticas — IDAL, Universitat de València

Verano 2026

Valencia, España · Próximamente

- Seleccionado para una estancia de investigación en el Intelligent Data Analysis Laboratory, el laboratorio de Machine Learning de la Universitat de València.

PROYECTOS

Sistema de Detección de Fraude — Proyecto Personal

Python · XGBoost · SHAP · FastAPI

- Pipeline end-to-end de detección de fraude sobre el dataset PaySim (Kaggle, ~6 M transacciones de mobile money), abordando desbalanceo extremo de clases.
- Comparativa de estrategias de sampling y evaluación mediante precision-recall AUC.
- Explicabilidad con SHAP y despliegue como API REST con FastAPI sobre Docker. Recall 99,76 % en el conjunto de test.

Iceberg Health — Plataforma de Datos Sanitarios

Python · FastAPI · Docker · AWS

- Plataforma de datos para sistemas de gestión de residencias geriátricas, desplegada en producción sobre AWS.
- Actualmente en fase piloto bajo NDA con clientes.

Iceberg Law — Procesamiento de Documentos Legales

Python · LLMs · NLP

- Extracción y estructuración automatizada de documentos legales y facturas mediante Large Language Models.
- Motor de anonimización GDPR/LOPD de PDFs en español on-premise: regex con checksums (DNI, NIE, IBAN), NER con BSC-RoBERTa y LLM local (Qwen3 vía Ollama).
- Pipeline reproducible con audit log SHA-256 y verificación post-redacción mediante OCR.

Iberia Geo-Alert — Proyecto Personal

Python · FastAPI · Sentinel Hub · Leaflet

- Dashboard real-time de eventos naturales agregando cuatro APIs públicas (USGS, NASA EONET, NASA FIRMS, Sentinel Hub).
- Análisis on-demand del Normalised Burn Ratio (NBR) sobre imágenes Sentinel-2 para incendios en la Península Ibérica.
- Backend modular FastAPI + Pydantic, frontend Leaflet con clustering.

FORMACIÓN

Grado en Ciencia de Datos — Universitat de València

2023 — 2027

Nota media 7,58/10 · Matrículas de Honor en Machine Learning y Probabilidad y Simulación.

Asignaturas: Machine Learning, Estadística, Bases de Datos, Series Temporales, Deep Learning.

Grado Superior en Informática — IES Abastos

2020 — 2022

Administración de Sistemas en Red. Prácticas de 4 meses en Slimbook: ensamblado de hardware y scripting en Bash para testeo automatizado.

PREMIOS

Mención de Honor — Concurso de Visualización de Datos de la Ciudad de Valencia

2025

Proyecto en equipo sobre zonas inundables de la Comunidad Valenciana, premiado por el Ayuntamiento de Valencia y la ETSE-UV en respuesta a la DANA.

COMPETENCIAS

Lenguajes y ML: Python, SQL, Machine Learning, NLP, Large Language Models.

Cloud y Herramientas: AWS, Docker, FastAPI, Bash, Git, Streamlit, Power BI.

Idiomas: Español (nativo), Valenciano/Catalán (nativo), Inglés (B2).