

FRANCISCO MINGUEZ

Independent Data Scientist & ML Engineer

Madrid, España | Remoto en Europa | Disponible para proyectos independientes

franminguezg23@gmail.com | [linkedin.com/in/francisco-minguez](https://www.linkedin.com/in/francisco-minguez) | franciscominguez.dev

RESUMEN PROFESIONAL

Data Scientist y ML Engineer independiente, basado en Madrid y disponible para proyectos remotos en Europa. Experiencia en el diseño y evaluación de sistemas RAG, modelos de machine learning y pipelines de datos con Python, PySpark y Databricks, junto con trabajos de calidad de datos y analítica de negocio. Trayectoria en tecnología, servicios financieros, seguros y regulación europea. Experiencia colaborando con equipos de producto, ingeniería y negocio, también en entornos con interlocución directa con clientes.

ÁREAS DE TRABAJO

- Diseño de sistemas RAG, evaluación del retrieval, generación fundamentada y análisis de fallos
- Machine learning aplicado, NLP y evaluación de modelos
- Pipelines de datos y ML con Python, PySpark y Databricks
- Evaluación de calidad de datos, analítica y automatización de reporting
- Colaboración externa con equipos europeos de datos, producto e ingeniería

EXPERIENCIA PROFESIONAL

Práctica independiente de datos e IA | Data Scientist & ML Engineer

Jul 2026 - Actualidad | Madrid / Remoto en Europa

- Desarrollo una práctica independiente centrada en analítica aplicada, ingeniería de datos, machine learning y sistemas de IA generativa fiables.
- Publico casos sobre fiabilidad RAG y modernización de pipelines analíticos en franciscominguez.dev.
- Desarrollo MigaDigital, una iniciativa paralela para pilotos acotados de datos, IA y automatización de procesos dirigidos a pymes españolas.

Multiverse Computing | ML Engineer

Ene 2026 - Jun 2026 | Madrid, España | Contrato por proyecto de duración determinada

- Diseñé y evalué pipelines RAG, con foco en calidad de retrieval, suficiencia de contexto, abstención y fiabilidad de respuesta.
- Desarrollé flujos de evaluación para LLMs y modelos vision-language, y trabajé con fine-tuning supervisado, compresión de modelos y recuperación de rendimiento posterior a la compresión bajo obligaciones de confidencialidad.
- Asumí la responsabilidad funcional de Product Owner en un workstream técnico, coordinando prioridades y criterios de aceptación.
- Presenté avances técnicos y resultados en sesiones con clientes y stakeholders, colaborando con equipos de ingeniería, producto e investigación.

Experian | Data Scientist

Sep 2024 - Ene 2026 | Madrid, España

- Desarrollé flujos de ML supervisado y no supervisado para casos de gestión de datos, segmentación y automatización.
- Migré y paralelicé un flujo de estimación de renta de SQL Server a Databricks/PySpark, mejorando su ejecución y mantenibilidad mediante componentes Spark reutilizables.
- Construí componentes PySpark reutilizables con controles automáticos de calidad de datos y validación estadística.
- Optimicé procesos ETL con PySpark y SQL Server/SSIS, mejorando la eficiencia de ejecución y la mantenibilidad.
- Desarrollé flujos de segmentación con K-Means y DBSCAN, utilizando PCA para reducción de dimensionalidad y análisis.

EXPERIENCIA PROFESIONAL

EIOPA - European Insurance and Occupational Pensions Authority | Data Scientist Trainee

Feb 2024 - Sep 2024 | Fráncfort, Alemania

- Construí un pipeline de NLP y OCR con spaCy y componentes basados en reglas para apoyar la revisión y priorización de documentación ESG.
- Medí el recall y analicé el efecto del flujo sobre el proceso manual de revisión documental.
- Desarrollé modelos predictivos con LightGBM y PyTorch y los evalué frente a un baseline establecido.
- Migré procesos analíticos basados en SAS a Python y Azure, mejorando su trazabilidad y mantenibilidad.
- Automaticé flujos de análisis estadístico y detección de anomalías con Isolation Forest y autoencoders.

AXA Seguros | Prácticas de Data Analyst

Sep 2023 - Ene 2024 | Madrid, España

- Migré procesos analíticos de SAS a PySpark y Databricks.
- Estandaricé métricas de negocio e introduje procesos analíticos más consistentes.
- Desarrollé análisis exploratorios y visualizaciones de KPIs con Python, R y Power BI para equipos de negocio.

Banco Sabadell | Director de Negocios / Experto en Seguros

Feb 2021 - Ene 2023 | Madrid, España

- Gestioné una cartera de clientes corporativos en productos bancarios y aseguradores.
- Utilicé segmentación, reporting y análisis de clientes para apoyar la captación, el seguimiento y la actividad comercial.
- Automaticé informes recurrentes y desarrollé experiencia en comunicación con clientes, negociación y productos financieros regulados.

PORTFOLIO SELECCIONADO

Modernización de pipeline analítico: Caso sobre la migración de procesamiento analítico legacy a flujos PySpark modulares con controles de calidad, logging, tests y análisis de rendimiento.

Estudio de fiabilidad RAG: Estudio sintético sobre calidad de retrieval, citas, abstención y análisis de fallos en sistemas de preguntas sobre documentos.

Casos publicados: franciscominguez.dev

COMPETENCIAS TÉCNICAS

Programación y datos: Python, SQL, PySpark, R

ML e IA: scikit-learn, LightGBM, PyTorch, NLP, RAG, evaluación de LLMs y VLMs, embeddings, OCR, compresión de modelos

Ingeniería de datos: Databricks, Spark, ETL, SQL Server, SSIS, controles de calidad de datos

Analítica y BI: Power BI, Streamlit, Plotly, diseño de KPIs, reporting analítico

Cloud y entrega: Azure, Docker, Git, GitLab

Colaboración: Presentaciones a clientes, comunicación con stakeholders, Product Owner funcional

FORMACIÓN

Máster en Data Science & Business Analytics, con doble especialización en Data Science y Business Intelligence: IMF / UCAV / Indra, 2023 - 2024. TFM: predicción de depósitos bancarios mediante modelos clasificatorios y una red neuronal polinómica profunda en PyTorch.

MBA: ENEB, 2020 - 2021

Grado en Marketing, impartido en inglés: Universidad Rey Juan Carlos, 2015 - 2020

IDIOMAS, CERTIFICACIONES Y DISPONIBILIDAD

Idiomas: Español - Nativo | Inglés - C1, TOEIC Listening & Reading: 910/990

Certificaciones: Docker, OpenWebinars (2024) | Scrum / Metodologías Ágiles, IMF (2023) | Fundamentos de Análisis de Datos, Microsoft (2023)

Disponibilidad: Inmediata | Remoto en Europa | Proyectos independientes | Contratos temporales | Colaboración con consultoras