FEDERICO FIORIO
Data & Machine Learning Engineer

Milan, Italy
+39 342 596 4621 |fioriofederico99@gmail.com
[GitHub](#) | [LinkedIn](#) | [Portfolio](#) | [Leetcode](#)

Results-driven **Data & Machine Learning Engineer** with expertise in **cloud-based data engineering, AI, and big data**. Proven track record in designing and optimizing **high-performance data architectures** at scale. Passionate about **building resilient, scalable data solutions and deploying ML models** to drive business impact.

EXPERIENCE

Data & Machine Learning Engineer | Data Reply, Milan, Italy
May 2024 - Present

- **Optimized Data Orchestration:** Migrated **Airflow DAGs** from legacy composer to **GCP Composer 2.x**, boosting **scheduling efficiency by 10x** for a global manufacturing client.
- **Enterprise-Scale Data Pipelines & Historical Data Optimization:** Designed **high-volume data ingestion pipelines** in **Databricks**, implementing **Medallion Architecture** for a fashion **luxury brand**. Replaced inefficient **JOIN-heavy historical tracking** with **SCD Type 2**, reducing query complexity and processing time by 100**%**, improving scalability, performance, and maintainability.
- **AI-Driven Document Retrieval**: Developed **RAG** pipelines for a leading bank, leveraging **vectorDB** and **OpenAI** embeddings, and custom document techniques to improve document search accuracy by 20%.

Data Engineer | Management Solutions, Milan, Italy
October 2023 - April 2024

- **IBM DataStage ETL optimization**: Developed ETL workflows with **IBM DataStage** and **SQL** for banking data warehouses, improving legacy **SQL** query by an average of 20%
- **Data Governance & Compliance**: Enforced **data governance** policies in cross-functional teams

Software Engineer Intern| Dilium, **Milan, Italy**
February 2021 - May 2021

- **GANs Reasearch**: Conducted research on GANs for deepfake generation, emphasizing ethical considerations
- **Training & Awareness**: Led training sessions for employees on the emerging risks of these technologies, especially audio forgery

EDUCATION

MSc in Computer Science | University of Milan
October 2021 – October 2023
Grade: 108 / 110

Erasmus+ **| University of Copenhagen**

BSc in Computer Science | University of Milan
September 2018 – September 2021

CERTIFICATIONS & PROJECTS

[Google Cloud Professional Machine Learning Engineer](#), [AWS Machine learning Specialty](#)

BoardGames RAG | [Repo](#)
This app leverages AI to make board game rulebooks interactive and easily searchable, allowing users to upload PDFs and get instant clarifications. It also offers features like saved chats and game references, all while providing a smooth, Docker-based experience.

Thesis: Membership inference attacks on text-to-image models | [Repo](#)
The aim of the Thesis is to test Diffusion Generative Models against Membership Inference Attacks and retrieve sensitive data

Competitive programming | [training & contests](#)

SKILLS

Python, Data Engineering, Machine Learning, Artificial Intelligence, Data structures & algorithms, Azure, Databricks, GCP, Composer, BigQuery, IAM, Airflow, AWS, API, Cloud Deployment, Databricks, Apache Spark, SQL, NoSQL, RDBMS, Scikit-learn, GitHub, IBM DataStage, VectorDB, Qdrant, Streamlit, Langchain, Docker.