

LUIZ FERNANDO COSTA DOS SANTOS

Data Engineering / Machine Learning Engineering

🌐 website 📞 +55 (83) 98691-5219 ✉ luizfernando1012000@gmail.com 💼 luiz-fernando632 🌐 luiz826

Languages: Portuguese (Native), English (Professional Fluency), Spanish (Basic)

EXPERIENCE

Nubank 🌐

Machine Learning Engineering

09 2022 – Present

Hybrid - São Paulo, Brasil

- Architected and refined large-scale Spark Scala pipelines, **reducing data processing time by 30%** and boosting model performance by **increasing AUC by 4 percentage points**.
- Led Python-based model orchestration in both batch and real-time production environments, achieving seamless integration within the company's data ecosystem.
- Migrated legacy ETL processes to a Data Mesh architecture, **reducing annual data infrastructure costs by over 25%** and improving data governance across **10+ product teams**.
- Ensured compliance with internal data governance and external regulatory requirements (similar to SOX), maintaining end-to-end data lineage and auditability.
- Partnered with product managers, data scientists, and business stakeholders to translate business goals into technical requirements, driving the successful adoption of new data pipelines.

Huawei and Embrapii 🌐

Machine Learning Engineering

02 2022 – 09 2022

João Pessoa, Brasil

- Developed a solar characterization dataset for predicting optimal solar tracker angles, leveraging real-world data from multiple Brazilian cities.
- Built an end-to-end ETL pipeline with Docker, Airflow, and InfluxDB to generate forecasts using a Random Forest model, **improving energy generation by up to 12%** under cloudy conditions.
- Enhanced data pipeline reliability by implementing automated monitoring and alerting for time-series forecasts, ensuring consistent performance and accuracy.

EDUCATION

Universidade Federal da Paraíba 🌐

Data Science and Artificial Intelligence

05 2025

GPA: 9.46 / 10.0

PROJECTS

PUMLE 🌐 |

2025

- A comprehensive data engineering pipeline for geological carbon storage applications, utilizing Spark to process and model geological survey data from Brazilian reservoirs.

Car Plate Detection 🌐 |

2024

- A computer vision solution built with Python and YOLO to identify and extract characters from Brazilian car plates, **achieving over 95% accuracy** on a custom dataset.

TECHNICAL SKILLS

Languages: Scala, Python, SQL, R, Java, C/C++

Data & MLOps: Spark, Databricks, AWS, Kubernetes, Docker, Airflow, Datomic, InfluxDB, Grafana, Postgres, Terraform, CI/CD

Frameworks: PyTorch, TensorFlow, Scikit-learn, Pandas, NumPy

Concepts: MLOps, Data Engineering, ETL/ELT, Data Mesh, Data Warehousing, Model Deployment, Distributed Systems, SQL/NoSQL

EXTRACURRICULAR

Momento 🌐

Mentor & Tech Lead

07 2025 – Present

João Pessoa, Brazil

- Mentor and guide team leads for the technology and communications fronts within a movement preparing and connecting university students to impactful careers in the tech industry.

TriLha 🌐

Advisor & Mentor

06 2024 – Present

João Pessoa, Brazil

- Served as an advisor and mentor, introducing new students to computer science via lectures, hands-on exercises, and project guidance. The project has impacted over 50 students across Brazil.