

# Leandro Fernandes

Current Location: Dallas, TX, USA

Email: leandrojfernandes@gmail.com

LinkedIn: [MyProfile](#)

## Introduction:

Senior Data Engineer with 6+ years of experience designing distributed data pipelines and integrating Machine Learning models at enterprise scale. Expert in Python, PySpark, and SQL optimization with a Bachelor's and a Post-doctorate in Systems Engineering. Proven track record of developing NLP solutions for major telecom clients (Verizon) and optimizing high-volume data architecture. Passionate about building robust, fault-tolerant systems that power intelligent features.

## Experience:

### Senior Data Engineer @ Mobileum (Dallas, Texas) | Feb 2022 - Present

- **NLP System Design:** Designed and implemented a production-grade NLP analytics engine to process semantic data from 25k daily targets. Engineered the end-to-end pipeline from model training to deployment, resulting in ~\$700k recurring revenue.
- **Infrastructure Optimization:** Architected distributed processing jobs using PySpark, optimizing compute resources to handle massive scale. Reduced query latency and resource consumption by 25% through algorithmic improvements and storage optimization (Trino/Hive).
- **Microservices & Serving:** Built and deployed low-latency inference services using FastAPI and Docker, enabling real-time consumption of ML models for client-facing applications.
- **Data Integrity:** Implemented robust validation layers using Couchbase (NoSQL) to ensure data consistency across distributed storage systems, supporting high-availability requirements.

### Application Data Engineer @ Mobileum (Braga, Portugal) | Jun 2018 - Jan 2022

- **Open Source Development:** Contributed core code to [nAttrMon](#) (an open-source monitoring tool), developing custom plugins to visualize real-time system bottlenecks. This enhanced observability across distributed clusters and reduced mean-time-to-resolution (MTTR) for outages.
- **System Performance:** Performed deep-dive performance tuning of Java applications (JVM) and Unix OS kernel parameters to meet strict latency and throughput SLAs in a 24x7 telecom environment.
- **Reliability Engineering:** Developed automated monitoring and self-healing scripts (Python/Bash) to minimize downtime, directly contributing to system stability for mission-critical carrier networks.

## Skills:

- **Languages:** Python (Advanced), SQL, Bash, Java.
- **Big Data & Storage:** PySpark, Hadoop, Networking (Certified).
- **Machine Learning:** NLP, Model Fine-tuning.
- **Databases:** ClickHouse (Certified), Trino, Hive, Couchbase (NoSQL), PostgreSQL.
- **Core competencies:** Data Pipeline Design, Performance Tuning, Distributed Systems, CI/CD.
- **Tools & DevOps:** Docker, FastAPI, Git, Unix/Linux.