

# MIGUEL FERREIRA DA SILVA

Lisbon | [miguel.marinho.silva@outlook.com](mailto:miguel.marinho.silva@outlook.com) | +351 933 852 719 | [LinkedIn](#) | [mmarinhofs.vercel.app](https://mmarinhofs.vercel.app)

## WORK EXPERIENCE

### NOKIA, Lisbon – Portugal

#### Data & Analytics - Data Scientist Lisbon – Portugal

Sept 2023 – Current

##### *Data Semantics Team*

- In charge of creating & enhancing complex supply chain data business objects (SQL) that can be used by business to facilitate end-to-end reporting, transforming their business requests to data engineering solutions
- Developed an advanced text generation Large Language Model (LLM), integrating metadata from data objects for facilitated information retrieval. Coded in Databricks workspace, mainly using MLflow, transformers.
- ECHO Δ: Scaled File Comparator to compare tables across multiple servers, featuring a user interface for customizable SQL queries.

Significantly streamlined the User Acceptance Testing (UAT) process by providing precise mismatch diagnostics in cloud migration.

Integrated with Azure DevOps for Continuous Integration and Continuous Deployment (CI/CD)

#### Data & Analytics - Data Science Intern Lisbon – Portugal

Sept 2022 – Aug 2023

Graduate program designed to rotate participants through key teams in Global Business Services - Data & Analytics

##### *Supply Chain Advanced Analytics Team*

- Classification project using Azure Cloud Machine Learning studio: Predicting On-Time Delivery of Supply Chain Products
- Migrated data pipelines from on-premises DB to Azure SQL DB; implemented solutions using Azure Data Factory

##### *Process Mining Team*

- Automating daily data extraction and API updates from Celonis API using Azure Function Apps
- Automated extraction, processing, and ingestion of Celonis audit logs API data into data pools using Python. Developed star schema, associated tables (SQL), & dashboard.
- ML Classification project using Celonis Machine Learning Workbench: Predict discount classification

##### *Data Semantics Team*

- Contributed on data migration tasks from on-premises server to Azure cloud, doing User Acceptance Testing (UAT) with SQL.
- Lead in developing a Python-based Excel File Comparator program, ensuring data consistency between on-premises servers and Azure cloud by precisely locating and reporting any discrepancies.

##### *Financial Planning and Reporting Analytics*

- Lead in optimizing SAP report generation by recreating complex reports using advanced SQL queries on a centralized data platform with ingested SAP tables, significantly reducing processing time.
- Python scripts for automated validation of reports.

## EDUCATION

Católica Lisbon School of Business and Economics, Lisbon – Portugal

Sept 2021 – July 2023

*Masters in Business Analytics – Specialization in **Data Science*** - GPA: 17/20

- Python (Pandas, Matplotlib, NumPy, SeaBorn, Git), Data Base Management (**SQL**), Big Data Technologies (**Hadoop, Spark, Hive**)
- Machine / Deep Learning: Executed a customer churn classification & Conducted an image classification project employing Convolutional Neural Networks (CNNs) with **TensorFlow**.

**Thesis: Deep Learning for Melanoma Classification: A Study Using Skin Lesion Images** - 19/20

- Set up & deployed multiple **Azure** servers for parallel training of MAR-MELA-CNN Fusion Ensemble of custom merged MAR-MELA dataset containing 15000 Melanoma images.
- Introduced **F $\beta$  Score** metric for medical applications to balance the trade-off between precision and recall, prioritizing the minimization of false negatives for critical diagnoses.
- **Scaled model for real-time inference**, integrated with Google Drive, to apply prediction on image uploaded from my phone and **deliver instant melanoma diagnosis via SMS**

Nova School of Business and Economics, Lisbon – Portugal

Sept 2017 - June 2021

*Bachelors in Economics*

European School of Luxembourg I

Sept 2005 - June 2017

*European Baccalaureate -*

Multilingual education in Portuguese and English, French as complementary language.

## SKILLS

- **Languages:** Portuguese (C2), English (C2), French (C1), Spanish (B1), German (A1- *Learning*)
- **Programming Languages:** Python (+PySpark), R, SQL
- **Machine Learning / Deep Learning:** scikit-learn, TensorFlow, Keras, transformers, huggingface
- **Data Visualization:** Tableau, Celonis Dashboards
- **Data Analysis:** Pandas, NumPy, SciPy, Matplotlib, SeaBorn
- **Data Engineering:** Git, Databricks, Azure Data Lake, Azure SQL DB, Azure Data Factory, ETL/ELT pipelines
- **Platforms:** Azure, Azure Machine learning Studio, Azure DevOps, SSMS, Oracle, Databricks, SAP S4/Hana
- **Certifications:** [AZ-900 : Azure Fundamentals Certificate](#) / [DP-100: Azure Data Scientist Associate](#) / [Celonis \(8\)](#) / [Finetuning Large Language Models](#) course DeepLearning.ai

## ACTIVITIES AND INTERESTS

- **Cultural multi diversity: Open-minded** & passionate in getting to know other realities and cultures. I love travelling, highlighting my trip to Asia.
- **Technology:** Highly connected to this sector from a young age. **Self-taught** assembly of a computer from scratch on my own, with over 40 hours of research to obtain highest performance at smallest market price.
- **Boxing:** Develops my **resilience**, my **determination**, **discipline** and **ability to stay calm under pressure**, due to having to meet with the various restrictions that occur during times of competition. Goal to be national champion in this category in 2024. **10+ hours** of weekly training
- **Chess:** Develops **strategic** thinking, analytical skills, decision-making abilities, **prepositioning** against possible adverse scenarios.