



- ✉ anascacais@gmail.com
- 🐙 github.com/anascacais
- in /anasofiacarmo
- 👤 anascacais.github.io

## Education

**PhD in Biomedical Eng.**  
*Instituto Superior Técnico*  
2021–2025

**MSc (Int.) in Biomedical Eng.**  
*Instituto Superior Técnico*  
2015–2021

## Tech Stack

### Programming Languages & Tools

- Python
- PyPI
- Jupyter
- SQL
- Git
- R
- unittest

### Data Analysis & Visualization

- Numpy
- Pandas
- Plotly
- Seaborn
- Matplotlib

### Statistics & Machine Learning

- Scikit-learn
- PyTorch
- SciPy
- Statsmodels
- metafor

### Backend & Databases

- FastAPI
- RESTful API
- SMB
- MySQL

## Other Skills

### Soft Skills

- Critical thinking
- Communication
- Team-first mindset
- Initiative

### Languages

Portuguese: C2, English: C2,  
German: A2

# ANA SOFIA CARMO

## Data Scientist

### Summary

PhD-level data scientist who learns new stacks quickly. Built end-to-end reliable pipelines, from a data-capture system to a MySQL + FastAPI API and diverse time-series models, now used in 3 hospitals and 5 research studies. Equally comfortable working in cross-functional teams and presenting complex findings to non-tech audiences (10+ talks, 3MT Ulisboa awardee).

### Experience

Feb 2021  
–  
Present

#### Research Assistant on Information & Data Sciences *Instituto de Telecomunicações, Portugal*

- **Technical Lead, Collaborative R&D:** Drove cross-institutional engineering efforts, shipping 2 clinically deployed research tools and co-authoring 10+ papers (*IEEE T-BME, Frontiers, etc*).
- **Infrastructure for Data-Driven Research:** Ran 2 GDPR-compliant protocols and engineered the capture system operating in 3 hospitals, standardizing workflows across sites.
- **Wearable-Sensor Innovation:** Developed a low-cost wearable device that now supports 5 ongoing research projects.
- **ML Modelling and Evaluation:** Built logistic regression and LSTM models (*Scikit-learn, PyTorch*); leveraged periodicity analysis (*SciPy*) to predict seizure-risk windows.
- **Open-Source Community:** Co-authored 3+ open-source Python packages, including biosignal processing, streamlined model development, and a client API.
- **Quantitative Methods:** Applied hypothesis testing (*SciPy*) and mixed-effects meta-analysis (*metafor*) to benchmark device performance and algorithm accuracy.

Dec 2024  
–  
May 2025

#### Biomedical Data Scientist (Freelance, part-time) *Remote*

- Automated feature engineering pipeline for 26 patients (*Numpy, BioSPPy*); ran mixed-effects models (*Statsmodels*) showing group-level changes post-intervention.

Sep 2021  
–  
Nov 2024

#### Teaching Assistant (part-time) *Instituto Superior Técnico, Portugal*

- Delivered 80h+ of hands-on ML & instrumentation labs; mentored multiple capstone projects; built Jupyter notebooks.

### Project Highlights

May 2025  
–  
June 2025

#### PreEpiSeizures - Seizure Prediction Database *Project Owner & Lead Developer*

- Designed a *MySQL* relational schema and built a *FastAPI* layer exposing secure REST endpoints to 2,500h+ of data.

Mar 2021  
–  
Mar 2025

#### ScientISST - Open Biomedical Engineering *Executive Leader (evolved from Social Media Manager)*

- Set strategic priorities, supported Ops., and helped guide our volunteer community's push for open biomedical innovation.