

# Joana Sofia Baião Vieira

☎ +351 919 818 111 | ✉ joanabaiao99@gmail.com | 🌐 joanabaiao.github.io | 📱 joanabaiao | 🌐 joanasbvieira

## PROFESSIONAL EXPERIENCE

---

### Bosch

Remote

#### Data Scientist

Feb. 2023 - Present

- Developed a ML model to predict employee turnover using 5 years of historical HR data. Built the model in Python with scikit-learn and Keras, and used explainability methods to identify key factors.
- Designed an optimised shuttle routing system for 45,000 employees in Stuttgart and Munich using Google OR-Tools. Integrated existing public transport for greater efficiency.
- Developed web scraping tools to collect 10,000+ job postings and applied NLP techniques and deep neural networks to extract key skills and analyse role similarities.
- Applied statistical and deep learning time series forecasting models to analyse crucial HR metrics such as headcount, average age, and exit ratio across 100+ Bosch locations globally.
- Developed and managed multiple ETL pipelines using Airflow, Python, and SQL to process and anonymise HR data from multiple sources, reducing workload by 40 hours per month.
- Currently analysing employee reviews using LLMs for sentiment analysis and categorising feedback by topic.

### CISUC

Coimbra, Portugal

#### Machine Learning Engineer (Research Fellow)

Nov. 2020 - Oct. 2022

- Developed a Machine Learning algorithm that predicts the course of Multiple Sclerosis in a specific time window using real-world clinical data. The pipeline was implemented in Python using the Scikit Learn and Keras libraries.
- This study achieved a remarkable +100% increase in F1-score compared to the base study.
- Explored explainability methods (SHAP, LIME and PDPs) to analyse and clarify the underlying behaviour of the developed ML models, providing valuable insights for improved interpretation by clinicians and researchers.

### jeKnowledge (Junior Enterprise)

Coimbra, Portugal

#### CEO & Innovation Consultant

Apr 2019 - May 2022

- Led a Junior Enterprise with 55 students, developing and supervising 68 [projects](#) (development, data science and design).
- Increased internal and external growth and the quality of the work developed, with a 224.76% surge in business volume, a 70.83% increase in projects, and a 6.83% rise in client feedback.
- Coordinated tech-related events: bluesessions, AI Sunsets, [Shift APPens](#), [.works](#), [jeKnowledge Academy](#), and [Coimbra.blockchain](#).
- Project Manager of [jeK Magazine](#) 1st-4th editions and Porto Blockchain Conference 1st edition.

### CFisUC

Coimbra, Portugal

#### Algorithm Developer (Research Fellow)

July. 2020 - Oct. 2020

- Developed a Python algorithm designed to reconstruct 3D medical images from 2D projections of vascular networks. This algorithm autonomously computes crucial metrics from the generated 3D views, including the number of vessels, their diameter, and the number of bifurcations.

## EDUCATION

---

### University of Coimbra

Coimbra, Portugal

#### Integrated Master's in Biomedical Engineering

Sep. 2017 - Oct. 2022

- Bachelor's degree in Biomedical Engineering and Master's degree in Clinical Informatics and Bioinformatics
- Relevant coursework: Machine Learning, Databases, Medical Imaging, Signal Processing, and Medical informatics,

#### Specialisation Course in Computational Biology

July 2020 - Oct. 2020

- Relevant coursework: Biostatistics, Genomics, Proteomics, and Modeling

#### Online certifications:

- [Deep Learning Specialisation](#) June 2024
- [Natural Language Processing Specialisation](#) Mar. 2024
- [AI for Medicine Specialisation: Medical Diagnosis, Prognosis, and Treatment](#) Dec. 2023

## SKILLS AND TECHNOLOGIES

---

**Programming Languages:** Python, SQL, Matlab, R, Java, HTML/CSS

**Frameworks:** Pandas, NumPy, Matplotlib, TensorFlow, PyTorch, scikit-learn, Keras, OpenCV, NLTK, Django, Flask

**Tools:** Git, Docker, Jupyter Notebook, Airflow, Azure, PowerBI

**Languages:** Portuguese (Native), English (Fluent), German (Basic - Currently learning A1)