

Hiring software engineer as easy as calling a taxi.

## Jorge D. Data Scientist, Biomedical Engineer

### SUMMARY

- 2+ years of experience as a Data Scientist at QIAGEN (biotech) and at DELOITTE (consulting)
- 2 years experience as Data Engineer & ML Researcher at CEREBRIU (AI & Medical Imaging)
- BSc in Biomedical Engineering, MSc in Data Science

### SKILLS

<b>Main Technical Skills</b>	Python
<b>AI &amp; Machine Learning</b>	Machine Learning, NumPy, OpenCV, PyTorch, Scikit-learn, TensorFlow, Transformer models
<b>Programming Languages</b>	R
<b>Python Frameworks and Libraries</b>	NumPy, Pandas, PySpark, PyTorch, Scikit-learn, SciPy, TensorFlow
<b>R Frameworks and Libraries</b>	Shiny
<b>Data Analysis and Visualization Technologies</b>	ETL, Microsoft Power BI
<b>Databases &amp; Management Systems / ORM</b>	SQL
<b>Amazon Web Services</b>	AWS Boto3, AWS S3
<b>Deployment, CI/CD &amp; Administration</b>	CI/CD, Kubernetes
<b>Virtualization, Containers and Orchestration</b>	Docker
<b>Version Control</b>	Git
<b>Operating Systems</b>	Linux
<b>Other Technical Skills</b>	Analytic Skills, Computer Vision, LLM, Medical Imaging, MLOps, nibabel, nilearn, NLP, pyarrow, skimimage, tidyverse, XNAT RedBrick



## EXPERIENCE

### **MACHINE LEARNING RESEARCHER, CEREBRIU (AI & radiomics)**

Sep 2023 - Present

Within the ML team and in collaboration with academia for my MSc thesis, I focus on automating the segmentation & detection of Cerebral Microbleeds in MRI brain data through Deep Learning.

- Continuous Literature research
- Design and coordinate the creation of a new clinically relevant dataset of microbleeds in collaboration with radiologists.
- Refinement and preprocessing of public and private datasets Training of 3D segmentation and detection models
- Investigating: multi-task and transfer learning, augmentations 37.5 ECTS. Preliminary results: <https://shorturl.at/bHLSW>

### **DATA ENGINEER, CEREBRIU (AI & radiomics)**

Nov 2022 - Present

In the Data Management & Analytics team, my efforts are directed towards on processing and managing data to facilitate deep learning model training within the Research department.

- Managing internal database of MRI images, annotation maps, and medical reports on Amazon S3 and XNAT
- Improving the processing pipeline for ingesting new hospital data
- Deploying Large Language Models for automatic labeling, classification, NER, and anonymization of medical reports
- Leading technical aspects of multiple data annotation projects through data analysis, image processing, and Sdk interactions
- Coordinating closely with medical annotators and radiologists

### **DATA ANALYST, Deloitte Digital (consulting)**

Sep 2021 - Aug 2022

Technical lead for various marketing automation and digitalization projects across automotive, banking, and airline industries.

- Developing automation scripts for ETL pipelines, analytics, KPI calculations, and SFMC campaigns launching & Surveillance
- Predictive modeling for customer behavior Data reporting to clients with real-time dashboards & reports



## **JUNIOR DATA SCIENTIST, QIAGEN (biotech)**

Nov 2019 - Sep 2021

Within the Data Science team, handling qPCR & production data.

- Developed custom R-Shiny apps and deployed on Linux servers R Backend coding (R package, API interactions, processing)
- Instituted a Data Catalog and coding standards for the team
- Administering Linux servers for computation & data hosting
- Producing statistical data analysis for various departments
- Tuning a classification algorithm for new DiagCORE panels

Prior to the Data team, I worked for 9 months in the R&D department for production as a Process Engineer. Scripting and doing wet lab tasks.

## **BSC THESIS - COMPUTATIONAL NEUROSCIENCE, IDIBAPS (research)**

Feb 2020 - Jul 2020

Within Systems Neuroscience research group. The goal was to investigate using long-temporal-scale information by various Recurrent Neural Networks to understand decisionmaking biases observed in real experiments with mice.

Contributed to NeuroGym open-source package.

- Article preprint: <https://osf.io/preprints/psyarxiv/aqc9n>
- Thesis report: <https://shorturl.at/mITW0>

## **EDUCATION**

### **MASTER OF SCIENCE - DATA SCIENCE**

IT University of Copenhagen | Sep 2022 - Jun 2024 (120 ECTS)

Grade: 10.2 (average), 12 (mode), 7-step scale

Theoretical and applied foundations of:

- Algorithm Design & Programming
- Advanced Statistics & Calculus
- Advanced ML & NLP (RNNs, CNNs, Transformers, GANs, VAEs)
- Production DS (Data wrangling, Databases, DevOps)
- Computer Systems Performance
- Medical Image Analysis & Processing

Projects can be found here: <https://shorturl.at/xIL37>.



## **BACHELOR OF SCIENCE - BIOMEDICAL ENGINEERING**

University of Barcelona | Sep 2016 - Jul 2020 (240 ECTS)

Grade: 8.7 (average), 9 (mode), 0-10 point scale 11 Distinctions with Honours

ENGINEERING knowledge: Algebra · Calculus · Differential Equations · Physics · Biophysics · Computer Science · Statistics · Programming · Electronics · Signal Processing · Computational Modeling · [Bio]Materials · Robotics · Bioinformatics · ML BIOMEDICAL knowledge: Cell & Molecular Biology · Biotechnology [Bio]Physics · [Bio]Chemistry · Pharmacology · Human Anatomy & Physiology · Tissue Engineering · Medical Instrumentation · Clinical Engineering · Nanotechnology · Medical Imaging

## **EXCHANGE SEMESTER - BIOMEDICAL ENGINEERING**

Delft University of Technology (Netherlands) | Jan - Aug 2019

Taking 60 ECTS of courses from MSc in Biomedical Engineering

