# DANIEL JORGE DEUTSCH | Data Scientist

Paris, Fr +33 7 45 10 02 91 danieljorged1@gmail.com daniel-jorge-deutsch daniel-deutsch.github.io

## PROFESSIONAL SUMMARY

Methodical and dedicated data scientist with 1 year of experience researching and developing deep-learning projects. Keenly interested in cutting-edge advancements in AI research and technology, with a demonstrated ability to navigate through complex and technical fields. Proficient in Python, adept at writing clean and robust code to engineer, design, and implement scalable machine learning systems. Eager to leverage AI and coding expertise to drive innovation and progress.

## SKILLS

| | |
|---|---|
| **Programming** | Python, SQL, C/C++. |
| **Python Libraries** | PyTorch, Pandas, NumPy, PySpark, Hugging Face, TensorFlow, Scikit-Learn, Hydra. |
| **Technologies** | AWS, GCP, Docker, Git. |
| **AI Experiences** | Deep Learning, NLP, Computer Vision, Transformers, RAG, Generative AI. |
| **Languages** | English (fluent - C2), French (fluent - C1), Spanish (intermediary - B2), Portuguese (native). |

## PROFESSIONAL EXPERIENCES

**VO2 Group**, Data Scientist - (Paris, Fr) Sep 2023 - now

- Designed and deployed a transformer-based neural network capable of generating text-conditioned outfit recommendations, significantly improving user engagement and experience in the platform.
- Implemented a deep-learning algorithm to predict the compatibility of an outfit's garments, achieving an 8% increase in AUC over the previous model.
- Engineered a Retrieval Augmented Generation (RAG) framework to improve the responses of large language models in the field of Environmental, Social, and Governance data, achieving a significant 11% accuracy boost.
- Optimized data preprocessing and sampling pipelines of a computer vision project, reducing the model's epoch training time from 15 to 3 minutes.

**Gravite.io**, Data Scientist - (Paris, Fr) May 2023 - Sep 2023

- Architected an unsupervised topic modeling algorithm to categorize user feedback and identify the company's main pain points, contributing to a 30% increase in relevant and actionable insights derived from customer input.
- Utilized parallel processing techniques to optimize the mentioned algorithm, leading to an 80% reduction in processing time and improving efficiency and scalability for feedback analysis.

**Kaiko**, Data Scientist Intern - (Paris, Fr) Sep 2021 - Sep 2022

- Conducted sentiment analysis research on social media content to verify its correlation with market trends, proposing metrics for the development of a sentiment index.
- Implemented unsupervised learning techniques to identify bot accounts across social media platforms, enabling an in-depth analysis of their influence on the crypto market.
- Investigated the application of anomaly detection algorithms on live market data, paving the way for more robust risk management strategies and improved market surveillance capabilities.
- Engineered a scalable PySpark pipeline tailored for processing vast volumes of historical market data, resulting in an 80% reduction in processing time and enhanced data throughput.

**BTG Pactual**, Data Engineer Intern - (São Paulo, Br) Dec 2019 - Aug 2020

- Designed a scalable Airflow architecture to orchestrate the usage of virtual machines for script executions, enabling a 40% increase in task throughput while reducing infrastructure costs by 25%.
- Developed a serverless framework to monitor the ingestion, processing, and enhancement of trading data, ensuring real-time feedback on data integrity.

## EDUCATION

**Télécom Paris - Institut Polytechnique de Paris**, Engineering Diploma - Artificial Intelligence 2020 - 2023
Coursework: Machine Learning, Deep Learning, Advanced Statistics, Speech and Audio Processing, Optimization.

**Paris 1 Panthéon-Sorbonne**, Master 2 (M2) - Finance Technology Data 2021 - 2022
Coursework: Financial Econometrics, Applied Data Science, Financial Quantitative Methods.

**Escola Politécnica - Universidade de São Paulo**, Engineering Diploma - Control and Automation 2017 - 2022
Coursework: Linear Algebra, Calculus, Probabilities, Statistics, Control Systems.

## CERTIFICATES

**Structuring Machine Learning Projects**, DeepLearning.AI | Coursera (verify) 2024
**Improving Deep Neural Networks**, DeepLearning.AI | Coursera (verify) 2023
**Neural Networks and Deep Learning**, DeepLearning.AI | Coursera (verify) 2023
**Foundations of Project Management**, Google | Coursera (verify) 2022
**Technology and Innovation Strategy**, Mines ParisTech (verify) 2020
**Cryptofinance**, Fundação Getúlio Vargas (FGV) (verify) 2019

## CERTIFICATIONS

**Green IT & Digital Sustainability**, GreenIT.fr (verify) 2024

## SCHOLARSHIPS

**Brafitec**, provided by CDEFI and CAPES to the most performant brazilian engineering students in France. 2021 - 2022

## AWARDS

**Panthéon-Sorbonne Master Thesis Competition**, 1st place 2022
**BTG Pactual Code Challenge**, 1st place 2019
**Brazilian Robotics Olympiad**, silver medal 2016
**Kangaroo Mathematics Olympiad**, silver medal 2016

## PROJECTS

**Effects of Social Media Bots on the Crypto Market** (verify) 2022
Sentiment Analysis, Bot/Anomaly Detection, Deep Learning, BERT, VAE, Data Visualization, Hypothesis Testing.

**How User Information Drives Amazon's Products Recommendations** (verify) 2021
Machine Learning, Statistical Analysis, Data Visualization, E-commerce Recommendation Systems.

**Analysis of Economic Incentives of Fake Reviews for Parisian Restaurants on Yelp** (verify) 2021
Sentiment Analysis, Data Scraping, NLP, Statistical Analysis, Data Visualization.