

Diogo Rosário
diogofiliperosario00@gmail.com
<https://www.linkedin.com/in/diogo-rosario>

EDUCATION

Universidade do Minho, Guimarães, Portugal

10/2019 – 06/2022

Bachelor's Degree in Applied Statistics

Related Coursework: Database Management Systems and SQL, Structured Data Analytics, Unstructured Data Analytics, Pricing Analytics, Statistics in Business, Data Analysis and Visualization, Time Series Forecasting, Statistical Learning, etc.

Universidade do Minho, Braga, Portugal

09/2022-

Master's Degree in Mathematics and Computer Science

Specialization: Machine Learning

PROFESSIONAL EXPERIENCE

Machine Learning Engineer – Issuu

May 2024 - Present

Worked on enhancing the company's machine learning systems by implementing comprehensive improvements in model evaluation. This involved developing and integrating robust evaluation metrics, conducting extensive comparative analyses of multiple models, and utilizing advanced validation techniques to ensure optimal model performance. Additionally, focused on improving the multilingual performance of one of the models by leveraging state-of-the-art natural language processing techniques. These efforts led to a more precise assessment of model accuracy, reliability, and generalization capabilities across different languages, enabling more data-driven decision-making and fostering continuous system improvement.

The technologies and programming languages used were: Python, PyTorch, Transformers, Hugging Face, Google BigQuery, and MySQL

Data Scientist Student Worker – Issuu

February 2023 – May 2024

Actively contributed to enhancing our machine learning capabilities, balancing responsibilities across labeling tasks, developing key machine learning projects, and supporting data analytics initiatives. Demonstrated proficiency in various technologies and programming languages to develop and implement machine learning models.

The technologies and programming languages used were: Python, PyTorch, Transformers, Hugging Face, Google BigQuery, and MySQL

Data Labeler Intern – Issuu

November 2022 – February 2023

Worked on Label Studio, labeling PDF files to feed to the company's Machine Learning models. The technologies used were Label Studio and Agile (Scrum).

Data Science Internship - Natixis

March 2022 – June 2022

Worked on simplifying Natixis' Market Risk Framework using Linear PCA, Rolling Window PCA, and Sparse PCA. Performed hyperparameter tuning to find the optimal parameters for Sparse PCA. The technologies and programming languages used were: Python, R, and scikit-learn.

Data Science Internship - Anybrain

July 2021 – September 2021

Worked on reducing the number of false negatives in the company's Machine Learning models. To accomplish this task, I did extensive data pre-processing, performed Hypothesis Tests on the data, and applied the results to the models.

The programming languages and technologies used were: R and h2o

SKILLS

Python, R, MySQL, Google BigQuery, JavaScript, React, HTML and CSS, Excel, TensorFlow, Amazon Sagemaker, Transformers, Hugging Face and Pytorch