ANDRÉ MARTINS

+(351) 911 053 772 ⋄ Guimarães, Braga

andrepizarromartins@gmail.com ⋄ GitHub ⋄ LinkedIn

PURPOSE

Currently working on my Data Engineering Thesis and looking for work experience in the area in order to learn more about it and evolve professionally and as a person.

EDUCATION

Masters in Software Engineering and Data Science, Universidade Do Minho	2022 - Expected 2024
Data Engineering Professional Certification, IBM and Coursera	2024
Erasmus, Eotvos Lorand University	2023
Big Data and Analytics, CISCO Networking Academy	2023
Bachelor of Computer Science and Software Engineering, Universidade do Minho	2019 - 2022

SKILLS

Technical Skills	Data Engineering, Data Analysis and Software Development
Soft Skills	Communication, Time Management, Problem Solving, Adaptability
Languages	Portuguese(Native) , English(C1), French(A2), Spanish(A2)
Programming Languages	Python, Java, SQL, HTML, CSS, C++

EXPERIENCE

Data Labeler and Data Science Junior for the ML team   November 2022 - August 2023
Issuu ,Braga and Copenhagen

- Learned about team planning and project management.
- Data Mining to generate datasets for labeling and labeling of the extracted data using Label Studio.
- Simple Business Intelligence through analysis of Data Lakes in Google Cloud using BigQuery.
- Created Copying script in python to migrate data in AWS S3.
- Feature extraction from datasets in order to train and deploy simple ML models.

Office worker doing invoicing and taking part in customer service   Summers of 2017 - 2021
Ribaembal , Portugal

- Developed communication, sense of responsibility and teamwork skills.
- Learned about interacting with client companies and business daily management.

PROJECTS

Data Lakehouse Implementation (Thesis) Research paper on the Data Lakehouse topic and developing a methodology for implementing one, and test of this methodology in a practical case.

Impact of Inflation on Mental Health and use of Social Media worldwide Discovering, treating, merging and analyzing 3 big Datasets in order to extract conclusions on the matter

School Platform Platform for School Management with interfaces for teachers, directors and students using Java, SpringBoot, React and PostgresSQL

Big Files C project reading data from different files with more than 3 Million lines with intertwined data, using complex DataStructures such as HashTables with the *glib* library.

Dataset Analysis and Model Training Analysis of a Roads and Incidentes Dataset from city of Guimaraes, Data treatment and Model training in order to maximize the accuracy of the predictions of road incidents. All made in python using pandas, numpy and sklearn libraries.