José Pinto

Curriculum Vitae

Personal Information

Name	**José Pedro Ribeiro Ferreira Pinto**.
Born on	**20/7/1998**.

Professional experience

Nov. 2020 – Now **Data scientist**, *INESC TEC*, Porto, Portugal.
Collaboration with Efacec on the project *transformer4.0*. Development of machine learning solutions for power transformer problems. Created a subset selection approach leading to significant lifecycle cost savings and reduced environmental impact. Deployed a cluster computing communication framework for IOT sensor data and distributed machine learning. Developed a synthetic data generation system to tackle shortcomings in data quantity and quality availability. Divulged results in the form of published scientific articles and company wide meetings.

Jul. 2020 – Set. 2020 **Data engineer (Internship)**, *INESC TEC*, Porto, Portugal.
Development of a solution for classification of LIDAR point clouds. Successfully developed and deployed a point cloud data generation and optimization approach, a wide array of deep learning computer vision models and automated hyperparameter optimization tools

Jul. 2019 – Set. 2019 **Software engineer (Internship)**, *Health Data Pioneers GmbH*, Munich, Germany.
Development of a web application for form creation with a drag-and-drop interface for the company's flagship software. Successfully delivered an MVP in 3 weeks, with the remaining time being spent on iterative feature improvement and addition.

Education

2019–2021 **Masters in Data Science**, *Faculty of Sciences of the University of Porto (FCUP)*, Porto, Portugal, Final grade average: 18/20; Thesis: 19/20.
Thesis: Identification of Viable Dissolved Gas Analysis Subsets for Power Transformers.

2016–2019 **Bachelors in Computer Science**, *Faculty of Sciences of the University of Porto (FCUP)*, Porto, Portugal, Final grade average of 16/20.

Computer Knowledge

Programming languages and markup

Python, R, C, Java, Javascript, AMPL, Prolog, HTML, CSS and LaTeX

Data science and Databases

Tensorflow/Keras, Pandas, Numpy, SPARK, Google Cloud, PostgreSQL, TimeScaleDB, MongoDB, MariaDB and Google BigQuery

Others

RabbitMQ, Kafka, Git/GitHub and Docker

Publications

Article **Optimal Gas Subset Selection for Dissolved Gas Analysis in Power Transformers**.
José Pinto et al. Under peer-review for the *International Journal of Prognostics and Health Management*.

Article **A flexible synthetic data generation framework for tabular data**.
Pinto and Sousa. Under peer-review for the *Engineering Applications of Artificial Intelligence*.

Languages

Portuguese	Native speaker
English	Full professional fluency