Vagner Zeizer Carvalho Paes

Brazilian
São José dos Pinhais /Paraná - (41) 99890-6987 - vzeizer@gmail.com – github.com/vzeizer
https://www.linkedin.com/in/vagner-zeizer-carvalho-paes - https://medium.com/@vagnerzeizer

Role: <div style="text-align:center">**Data Scientist**</div>

Summary:

Data Scientist with **4 years** of experience **and 11 years of academic experience**, who completed three post-doctorates in Computational Physics. Nowadays, Vagner is a member of the PUC-Rio ECOA Institute team responsible for supporting the development of the **IBRFlex** Project, in partnership with **Petrobras**, in a hybrid role of Data Scientist and Full-Stack Web Developer.

Vagner worked as a Data Scientist at the Pontifical University of Rio de Janeiro (PUC-Rio) in a **R,D&I** (Research, Development and Innovation) project named FURNAS/Ensights. In this project, Vagner has been engaged with application of data mining, data cleaning/wrangling and development of machine learning techniques for the intelligent monitoring of transmission assets. In his previous and first experience in a Data Science Role, he optimized the performance of medical treatments by using state-of-art Physics, Medicine and Applied Machine Learning.

He has solid experience in computer programming in languages such as *Python 3*, *C++, and SQL* by using libraries and frameworks such as *NumPy, Pandas, Seaborn, PyTorch, TensorFlow, and Scikit-learn*. He also has real-world experience with software development, by using technologies such as Docker, Django, React, JavaScript, HTML and CSS.

He has published **20 articles** in international journals, such as *Physical Review B, Journal of Magnetism and Magnetic Materials, RSC Advances* and *Journal of Physical Chemistry C*, contributing to modeling of nanostructures and obtaining mastery in writing and presentation in English. Out of these 20 articles, **4 were published in Data Science related journals**, such as, **IEEE**, **Research, Society and Development** and **Journal of Health Informatics**. For more details, see my Lattes at http://lattes.cnpq.br/8017238644593787.

He also has intermediate knowledge of Azure and AWS Cloud Services.

Languages: Fluent in English. German and French intermediate proficiency level.

Professional Path

Instituto ECOA (PUC-Rio)
Role: Data Scientist/Full-Stack Web Developer

Main projects: **07/2024-onwards**
- The project involves implementing a system with various integrated computational tools that, based on the methodology and development paradigms established by TIC/Petrobras, allows editing of input data and engineering rules, collects data from reditors, issues diagnostics with good visual presentation, suggests resource optimizations, enables what-if simulations, and supports risk-based inspection decisions for all flexible pipelines in operation in Petrobras' subsea assets.

Main Activities
- developing applications by using Python (framework Django), JavaScript (framework React/Next.js), HTML and CSS. The apps are built by using docker.
- Performing statistical analysis by using Python programming language and making them available to authorized users;
- aided at the development of scientific applications according to Monte Carlo simulation procedure, making UIs also;
- helping write project documentation in order to assist the software development lifecycle;

PUC-Rio   **04/2023 – 06/2024**
Role: Data Scientist

Main Projects

- Participated in the development of the whole data science project to create models for predicting defects in power transformers through Machine and Deep Learning techniques. By using analogical and sensor data from power transformers collected over a period of 8 years, the main goal of the project was to create indicators that could potentially predict the occurrence of failures in power transformers with some days in advance, saving millions of dollars of the company annually.

Main Activities

- assisting and developing high quality code in Python for data cleaning, structuring, as well as integrating scripts from different sources;
- assisting and developing Machine and Deep Learning models to predict defects in power transformers using classification models, as well as anomaly detection models. The models achieved an AUC of about 0.9, predicting the occurrence of a defect with some days in advance, potentially saving millions of dollars of the company annually;
- handling source code in a repository through Azure DevOps;
- participated in the writing and presentation of scientific articles concerning the main achievements of the project;
- prepared several deliverables concerning documentations of the project, such as the models documentations and Data Quality Reports;
- assisted preparing machine learning models for production, putting them in customized pipelines.

Bright Photomedicine   **05/2021 up to 03/2023**
Role: Data Scientist

Main Projects
- Found out the days interval in which a patient needs to do the treatment for a determined pathology aiming at obtaining betterment of the patient's pain after the ending of the treatment.
- Found out the main patterns and analyzed the outcome of the patients along the treatment as a function of its attendance frenquency.
- Responsible for finding patterns and data analysis of a clinical trial getting valuable indicators about the efficiency and power of a *Chatbot* to monitor the evolution of the treatment of the patients. Obtained perspectives on how to improve the tool through data analysis. In only one month the data analysis was performed and synthesized in scientifc value for the startup.
- Presented the results of a *Chatbot* as a tool to monitor the evolution of the patients in an International Conference, named CBIS (Congresso Brasileiro de Informática em Saúde).

Main Activities
- Analysis of the evolution of the quality of life of patients along the treatment, while writing scripts and codes mainly in *Python 3*.
- Writing and publishing of articles in national and international journals (Research, Society and Development; Journal of Health Informatics), aiming at increasing the scientific basis of the treatment offered by the Startup, impacting the commercial results.
- Identifying the main features of the patients that influence in the outcome of the treatment in the long-term.
- Clustering of patients into groups according to, for instance, skin type and age range as a tool to find the profile of adhesion of the patients to the treatment.
- Elaboration of indicators to evaluate the evolution of the patients session by session, as well as in the beginning of the treatment (anamnesis), providing indicators of tendency of improvement or worsening along the following sessions of treatment.
- Data analysis of patients undergoing low-back pain, finding tendencies and indicators of improvement in quality of life. Delivered business value from the data and published in scientific journals (Research, Society and Development).
- Development of a documentation for the steps of the data analysis, data processing, and future machine learning model for improvement of the treatment. Elaboration of a document with all

the steps and procedures to evaluate the quality of life of patients in pilot studies and clinical trials.
- Data analysis of patients with applied dosis in cooperation with an enterprise. We obtained great results for the evolution of the patients along the treatment. Found out gaps in the treatment as a function of the frequency of the attendance of the patients, opening room for further improvement of the treatment.
- Investigation of data of patients of an ambulatory and found out a significant tendency of improvement in the patients' pain during four weeks of treatment.

UFRGS   **03/2020 up to 04/2021**
Researcher

Main Activities

- Optimization of Python 3 codes to describe the magnetization reversal through the minimization of the free energy density or the coupled LLG equations.
- Implementation and theoretical approaches to study interactions in magnetic systems.
- Application of the Monte Carlo method in *Python* and *C++* to study the magnetic behavior of exchange biased systems.

Universidade Federal do Paraná   **08/2018 up to 08/2019**
Researcher

Main Activities

- Broad domain of programming languages, such as *Python 3* and *C++,* obtaining several results in the branch of modeling of nanostructures, which were published in over 10 international journals.
- Mastering of *Python 3*, aiming at high performance through numba (JIT compiled code) and parallelization using *joblib*.
- Implementation of Monte Carlo methods in *Python* and *C++* to study the magnetic behavior of nanostructures.

Academic Background

- Post-Graduation Specialization in Data Science and Analytics – PUC-Rio – 2023 - 2024. Grade: 9.5
- Post-Doctorate – Simulation of Exchange-Biased Nanostructures – UFRGS - 2021
- Post-Doctorate – Simulation of Shape Memory Alloys - 2019
- Ph.D in Physics – Simulation of Catalytic Nanostructures - UFRGS - 2017
- Master in Physics – Analysis and Modeling of Magnetic nanostructures - Universidade Federal do Paraná - 2012
- Bachelor's Degree in Physics – Universidade Federal do Paraná - 2009

Main Specializations/MOOCS

- Deep Learning for AI applications with PyTorch and Lightining – Data Science Academy - 2024
- Image Analysis with Artificial Intelligence – Data Science Academy - 2024
- Statistical Analysis for Data Science with Python Language – Data Science Academy - 2023
- Time Series Analysis and Forecasting with AI – Data Science Academy - 2023
- Business Analytics – Data Science Academy - 2023
- Big Data Real-Time Analytics with Python and PySpark – Data Science Academy - 2023
- Time Series's Analysis and Forecasting with AI – Data Science Academy - 2023
- Cybersecurity Data Science – Data Science Academy – 2023
- Academy Accreditation - Databricks Lakehouse Fundamentals - Databricks - 2023;
- Nanodegree Machine Learning Engineer with Microsoft Azure – Udacity – 2023
- AI for Medicine - Coursera – 2022
- Deep Learning Specialization - Coursera – 2022
- Machine Learning Engineering for Production (MLOps) - Coursera – 2022

- Practical Data Science on AWS - Coursera – 2022
- Data Scientist Nanodegree – Udacity - 2021

Experience abroad

TU Chemnitz - Germany – 2019
- Approaches to model the magnetic behavior of shape memory alloys.
- Optimization of Python 3 codes to describe magnetic systems.

Awards

- Best Paper Award - On the use of Machine Learning for Predictive Maintenance of Power Transformers - Issued by XVI Brazilian Conference on Computational Intelligence - October 2023.