Felipe Glicério Gomes Marcelino

Data Scientist

& (+55) 31-98765-5482 | ⊠ felipe.marcelino1991@gmail.com | ♀ FelipeMarcelino | ♀ https://

felipemarcelino.github.io/

Summary

Data Scientist with over 6 years of experience, specializing in machine learning, statistical modeling, big data, data analysis, and software engineering. Proven ability to communicate effectively with teams and stakeholders, transforming business problems into AI solutions using data-driven insights and visualizations. Extensive experience in finance (fraud detection), industry 4.0 (steel quality prediction), healthcare (glaucoma detection via computer vision), resource allocation (reinforcement learning), retail (demand prediction and pricing), and energy (price optimization). Proficient in Python and Scala, with strong software engineering skills. Adaptable to different environments, excellent communicator, and team player. Holds a Master's degree in Artificial Intelligence and a Bachelor's degree in Computer Science.

Skills

Data Science: Machine Learning, Statistical Modelling, Data Visualization, Data Analysis, Research, Big Data, ETL Programming Languages: Python, Scala, Lua, Nix (NixOS)

Tools and Frameworks: Github, Pytorch, Scikit-Learn, Apache Spark, Pandas, IATEX **Languages:** English (proficient)

Experience

Data Scientist

Nubank

12/2021 - Present

- Develop project on KYC (Know you customer) that involves fraud identity, where the fraudsters try to be another person and obtain any credit product from Nubank. Using automatic retraining procedure and including new features to the model we had a improvement of 20% on precision and 15% on recall making the operation to catch fraudster less costly and less friction to enter on Nubank and legitimated customers to acquire products.
- I develop model and maintain to catch mule account models, using transactions from different types (PIX, Debit, Withdrawal and, etc.). A mule account involves using an account to hide the origin of the fraudulent money from any kind of criminal activity. Using graph features, the model improved 27% on the precision and, 31% on recall, making the operation cost to catch fraudster decreasing almost 37%.
- $\circ~$ Besides that, I participate on D&I task force to support minority inside Nubank.

Data Scientist

Kunumi

12/2020 - 12/2021

- Demand Prediction: The goal was to prevent products from going out-of-stock, ensuring they were always available for customers to purchase. Boosting models were used on time series data to determine, within a range, the optimal number of units (SKU) for a specific point of sale across the entire country.
- Price Elasticity: Using a regression model, the idea was to select a specific price that would yield the highest
 possible profit for the company through the sale of units of different SKUs. The price was chosen based on the line
 drawn by the regression. The resulted model increase the revenue in 18%
- Resource Allocation: Using reinforcement learning methods, the idea was to determine the best number and specific boats to be used in the docking of cargo ships. Fuel consumption varied for each boat due to different power levels. Additionally, as different cargo ships arrived at different ports, it was also necessary to use parallel localization and allocation. The resulted model decreases the cost to allocate the correct resource, plus the usage of fuel, decreasing on total of 4%.

Data Scientist/Research AssistantFederal University of Minas Gerais10/2018 - 12/2020

• APERAM Project:

- Develop and generate explanations for Machine Learning models in the steel manufacturing process.
- $\circ~$ Pre-process and prepare data for ML tasks.
- $\circ~$ Train ML models on steel manufacturing data to predict defects in steel sheets.
- Use explainable ML techniques to clarify model decisions, providing insights to engineering experts on how to improve steel recipes.
- The result of the project generates a paper cited on P1

• PSR Project:

- Develop Machine Learning models for the dispatch of electric power production from multiple sources.
- $\circ~$ Pre-process and prepare data for ML tasks.
- Train ML models using climate and structural data from energy production sources, such as hydrothermal plants, using reinforcement learning techniques.
- Create baselines for comparisons.
- Analyze results and consistency in collaboration with the partner.

Data Scientist InternTecsinapse06/2017 - 08/2018

• Nanoparticles:

• The overall goal of this project is to identify and develop a computational model, using neural networks, capable of predicting and simulating the green synthesis of metal nanoparticles based on real data generated by experiments conducted through the partnership.

• NLP:

• The overall goal of this project is to use natural language processing models, bi-LSTM and GRU, for sentiment analysis. The textual data provided by salespeople at a car dealership is processed by the model, and through sentiment analysis, the probability of a customer returning to finalize a purchase is determined.

IT Project Manager

Informática Junior

06/2016 - 08/2017

- $\circ~$ Managed web application development projects using Scrum.
- Supported directorate processes, including commercial proposals and project monitoring.
- Handled the commercial aspects of the organization, improving project closure efficiency.
- Responsible for supporting the execution of the directorate's processes, such as creating commercial proposals, researching the technical competencies of company members, and monitoring projects. Execution of projects using JavaScript (Node.js, Express), WordPress, and the MVC pattern.

Research Assistant

Federal University of Minas Gerais

11/2015 - 11/2016

• Analyzed developer profiles and productivity impacts of role transitions.

• Investigated tool usage patterns among developers.

Education

Federal University of Minas Gerais	Minas Gerais, Brazil
M.S. in Artificial Intelligence	01/2019 - 02/2023
B. in Computer Science	08/2013 - 08/2018

Publications

P1 Predicting Heating Sliver in Duplex Stainless Steels Manufacturing through Rashomon Sets.
 Gianlucca Zuin, Felipe Marcelino and, et al.
 International Joint Conference on Neural Networks (IJCNN) 2021