

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, \LaTeX

Languages: English (proficient)

Experience

Data Scientist	Nubank	12/2021 - Present
<ul style="list-style-type: none">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%.Besides that, I participate on D&I task force to support minority inside Nubank.		
Data Scientist	Kunumi	12/2020 - 12/2021
<ul style="list-style-type: none">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 Assistant	Federal University of Minas Gerais	10/2018 - 12/2020

