

PATRICK GOMES DE OLIVEIRA

Data Scientist

+5521976639291 - patrickufrj@gmail.com - Rio de Janeiro, RJ, Brazil

<https://github.com/Patotricks15> - <https://www.linkedin.com/in/patotricks15/>

SUMMARY

Data Scientist with 4 years of experience. My expertise focuses on developing Machine Learning models, Econometrics, Generative AI and Exploratory Data Analysis. Proficient in Python and SQL, with a solid foundation in databases like MySQL and MongoDB, LLM and Agentic workflow with Langchain and Langgraph as well as cloud platforms such as Google Cloud and with basic knowledge of AWS.

EDUCATION

Bachelor in Economics Science - UFRRJ

Offered mini-courses in Python and SQL

Research in Applied Economics with Econometrics and Machine Learning

Establishment of a Data Science league for academic projects and research

Project development in technology applying LLM (Vertex AI) to build chatbots to teach geology for kids

EXPERIENCE

Stefanini [client: Bank of Brazil]

Data Scientist

September 2023 - Present

São Paulo, SP

- **Exploratory Data Analysis:** Used statistical techniques like Test-T, correlation analysis and econometrics models with fixed-effects and others statistics to find patterns in financial product data;
- **Machine Learning:** Used Machine Learning techniques on time series for bank liquidity forecasting using monetary variables (reducing losses in about R\$600 Mi), avoiding overfitting, data shift and data leakage;
- **Machine Learning Engineering:** Application coding of routines for training and storing machine learning models in order to measure accuracy over time and facilitate model maintenance. With the template project, the team reduces about 60% in time.
- **Applied Technologies:** Scikit-Learn, Pytorch, Tensorflow, Python, Statsmodels, Pyspark, Dbeaver. In this company, pyspark is used in every project to capture a big amount of data to be used in machine learning models and data science projects.

ZapGPT

Data Scientist

February 2023 - September 2023

São Paulo, SP

- **Leadership:** Established the data science department within a early-stage company, building bridges with the Business, Product, Marketing, and Software Engineering teams. It developed a lot of my communication, ownership, leadership and collaboration skills;
- **Generative AI:** Backend develop to build AI Agents from documents submitted by users and extracted information, refining prompts and establishing best practices in the development of POC / MVP / products LLM-powered (Hugging Faces and LangChain).;
- **Machine Learning research:** Used Machine Learning techniques to estimate retention, churn, forecast ROI, and subscription probability, the frameworks used were TensorFlow and scikit-learn. My set models give the information 300% faster than a CEO's traditional Excel Sheet, with 80% less errors.
- **Applied Technologies:** LangChain, Hugging Faces, Python, Scikit-learn, Numpy, MySQL. Applied these technologies to develop Generative AI applications and machine learning models. Streamlit was used to build POCs and MVPs.

RankMyAPP

Data Scientist

September 2021 - January 2023

São Paulo, SP

- **Product building:** Developed an internal and external MVP, POC and tool for geolocated keyword search; developed two tools for use in exploratory data analysis, focusing on sentiment analysis (NLP) in Play Store comments. Both products were used to generate revenue to the company.
- **Generative AI:** Responsible for identifying bottlenecks in customer service, leading to the conceptualization of a product using LLM and NLP. This initiative included the ultra-specialized refinement of prompts to respond to Play Store comments, significantly reducing costs by about 90%.
- **Exploratory Data Analysis:** Created analyses and reports for the marketing team using statistical analysis, which were published on the company's website and blogs, generating visibility for the company.
- **Applied Technologies:** Python, Scikit-learn, Numpy, Matplotlib, NLTK, OpenAI API, GeoPandas, Plotly, Leaflet, MongoDB, Airflow. Applied machine learning models to datasets about mobile apps, conducted geospatial analysis, and built ETL pipelines in data engineering projects.

KeyCash

Data Scientist

February 2021 - August 2021

São Paulo, SP

- **Exploratory Data Analysis:** Conducted exploratory data analysis using Python, pandas, numpy and Jupyter Notebook to uncover insights, addressing critical business questions and supporting evidence-based decision-making in real estate financing.
- **Artificial Intelligence:** Implemented computer vision using Street View and Google Cloud Vision APIs for advanced pattern recognition in data, reducing the time in house photo analysis on 92
- **Data Engineering:** Created Python scripts for automated document collection and spreadsheet manipulation through web scraping and crawlers, significantly improving operational efficiency and accuracy in data handling.
- **Applied Technologies:** Python, Google Cloud Platform, Open Vision, MySQL, Docker, Airflow, requests, BeautifulSoup. Applied computer vision to recognize objects in house photos and built ETL pipelines for web scraping projects.

VOLUNTARY WORK

CNPQ

Researcher in Applied Economics

August 2021 - August 2022

Rio de Janeiro, RJ

- **Project:** Assessment of Poverty in the Brazilian Economy Using Microdata from Continuous PNAD and Machine Learning.
- **About:** investigation about social-economic problem using Logit model, and a machine learning model to predict the poor during the COVID-19 pandemics

CNPQ

Researcher in Technology

January 2023 - January 2024

Rio de Janeiro, RJ

- **Project:** Interactive Geology: Conscious Society.
- **About:** Led the data science team in establishing a comprehensive database of minerals used on the website, and the implementation of Generative AI technologies (LangChain and Google Vertex AI)

MAIN SKILLS

Programming Languages

Python, R, SQL

Databases

MySQL, MongoDB, VectorDatabases, Firebase, Spark, Pyspark

Knowledges

API, Spatial Data Analysis, Airflow, web scraping

Dashboards

Power BI, Grafana, AWS QuickSight, Looker Studio

Cloud Computing

Google Cloud Platform, Amazon Web Service

Soft Skills

Intrapreneurship, ownership, problem-solving, business metrics, creativity

Languages

English, Portuguese, Spanish