LUCAS MORAES

I am a data professional who transits between data science and analytics. I am an evolutionary biologist by formation, having acted with statistical modelling applied to the field of bioinformatics. I program in R and Python and while I am not developing models, I am analysing them to assure their quality and scientific rigor.

💼 RECENT PROFESSIONAL EXPERIENCE

Present | 2022

Senior Data Analyst
PicPay

- Statistical and experimental support for the User Knowledge Squad.
- Data analysis for the development of robust machine learning models.
- Data analysis for integrity check of models in production.

2022 | 2018

Independent consultant (Data Science & Analytics)
Freelancer

- Data compilation, cleaning, exploratory analysis and statistical modelling for reports and research projects.
- Experimental design and hypothesis testing for the resolution of questions with scientific, statistical and methodological rigor.
- Development of machine learning models for predictive or correlation analyses (e.g. linear and logistic regression, k-means, random forest and XGBoost).

2022 | 2021

Data Scientist
Melhor envio

- Customer segmentation using non supervised machine learning (K-prototypes).
- Supevised machine learning for customer churn prediction (Random Forest & XGBoost).
- Development of a analytical pipeline to monitor the activities of customers in real time, with the objective of increasing retention and detect churn, using personalized behavioral data.
- Conversion of arbitrary business metrics to robust indicators using statistical techniques (e.g. bootstrapping and hypothesis testing).
- Data viz e dashboards (ggplot2 & Looker). Presentations to non techincal audiences.

Contact & Portfolio

🏠 lucasmoraes.org

✉ moraessaur@gmail.com

🐙 GitHub: @moraessaur

in LinkedIn

🎓 EDUCATION

2018
|
2016

MsC, Genetics

Rio de Janeiro Federal University

- Hierarchical clustering and dendrogram analysis of dated phylogenetic trees to identify evolutionary distinct angiosperm lineages, integrating biological, geographical and molecular data.
- Dissertation: Conservation of evolutionary distinct brazilian angiosperm species: integrating extinction risk assessments, phylogenetic information and the state of the art knowledge of brazilian plants.
- Advisor: Carlos Guerra Schrago.

2012
|
2007

BsC, Genetics

Rio de Janeiro Federal University

- Phylogenetic and topological estimation of cetaceans using bayesian and maximum likelihood inferences for hierarchical clustering parametrization.
- Mithocondrial genome sequencing and in silico analysis.
- Monography: Phylogenetic Status and Timescale for the Diversification of Steno and Sotalia Dolphins
- Publication: Phylogenetic Status and Timescale for the Diversification of Steno and Sotalia Dolphins. PLOS ONE. https://doi.org/10.1371/journal.pone.0028297
- Advisor: Carlos Guerra Schrago.

Technical knowledge
——————————

Ⓡ R

🐍 Python

🗄 SQL

Spark

📊 Statistics

🤖 Machine Learning

📈 Data Viz

♀ Fluent english

👥 Soft Skills

📖 ABOUT ME

A few considerations

- I have been part of interdisciplinary teams with people from a wide variety of backgrounds and seniority levels. This gave me a keen sense of empathy and understanding about people in general. This is for me one of my most valuable assets.
- I believe in the principle of parsimony: the best way is the simplest possible, although the simplest may be complex.
- I love developing models, but I believe that well treated data coupled with a rigorous experimental design is way more important than modelling.
- This cv was generated in R!

Curiosities

- I have previously acted as a professional photographer.
- I am always on time.
- I practice free diving and spearfishing.
- I spent the beginning of my childhood in Wyoming.