LUCAS CARAMÊS
+55 91 98321 6206
lgpcarames@gmail.com

Manaus, AM,
Brazil
Linkedin
GitHub

Introduction

I am a Ph.D. in Statistical Physics from the Universidade Federal do Rio Grande do Norte, with extensive experience in statistical and computational modeling, focusing on the study of optimization algorithms. I have experience in developing Machine Learning models, insights, and analyses using statistical tools and data visualization. I have participated in constructing data pipelines, creating models, building APIs, and developing web applications. I am skilled in tools such as Apache Airflow, GCP, Looker Studio, GIT, Wolfram Mathematica, Matlab, Redash, Mongo, SQL, Python, C, Flask, Torch, Docker, JavaScript, HTML, CSS, among others.

Experience

2023-PRESENT **Data Scientist**
Development of web applications with integrated Artificial Intelligence (AI), offering generative AI solutions employing reinforcement learning models based on human feedback. Skilled in creating dashboards through Looker Studio, building recommendation systems, developing RESTful APIs, and deploying on GCP utilizing BigQuery, CloudSQL, Artifact Registry, GIT, and CloudRun.
Institudo de Pesquisas Eldorado. Manaus, AM - Brazil.

2022-2023 **Data Scientist**
Throughout my professional career, I have had the opportunity to work comprehensively across various domains, including data visualization, modeling, and data engineering. However, my primary focus has been on developing models and conducting studies aimed at business intelligence and decision support. I have specialized in developing web applications with integrated Artificial Intelligence (AI).
Gyra Mais Tecnologia. São Paulo, SP - Brazil.

Academic Trajectory

2014–2018 **BACHELOR'S DEGREE PROJECT:**
I investigated how fields originating from quantum vacuum fluctuations interact with matter depending on the boundary conditions imposed by the matter on the fields. For this work, I utilized various scientific computing tools such as Maple, Maxima, Wolfram Mathematica, and also gained experience with FORTRAN and Matlab throughout my undergraduate studies. LyX and LaTeX were my primary text editors.
Universidade Federal do Pará. Belém, Pará - Brazil.

2018–2019 **MASTER'S THESIS**
I developed a robust theoretical model, from first principles, describing the electromagnetic interactions of a graphene sheet with another nearby material, highly applicable in the study of new graphene devices and in designing experiments that explore relativistic phenomena in materials. During this period, I also utilized the same scientific computing tools employed during my undergraduate studies.
Universidade Federal do Pará. Belém, Pará - Brazil.

2019–2023 **PH.D. THESIS**

I hold a Ph.D. in the field of complex systems physics, where I specialized in researching optimal random search strategies aimed at explaining behaviors in biological systems, financial markets, and optimization algorithms. In addition to using the previously mentioned tools, I also routinely worked with Python and C languages for the development of the project.

Universidade Federal do Rio Grande do Norte. Natal, Rio Grande do Norte - Brazil

2023–PRESENT POSTDOCTORAL RESEARCHER

I lead a development group to modernize legacy codes of the Alpha Social Accounts, a methodology developed and widely used to refine the calculation of wealth generation from the flow of goods in a specific region. For this project, we utilize tools such as Python, Tox, GIT, and Travis for implementing CI/CD workflows.

Universidade Federal do Pará. Belém, Pará - Brazil.

Projects

DATA ENGINEERING PROCCESS

Throughout my professional experience, I have been involved in various projects concerning data cleaning, processing, and creation of new variables. This process often involved ETL workflows orchestrated using Airflow DAGs, or depending on the workload, using a Docker project along with Kubernetes, served via a Flask API.

DECISION-MAKING STUDIES

During my professional experience, I actively participated in studies to provide technical assessments to the board, aiding in crucial decision-making processes. Among these studies, I highlight my involvement in profiling customers based on banking variables and commerce sector, as well as leading various analyses on the impact of changes in the credit conveyor, business model, or other initiatives. For such studies, I always relied on the Python language, working with SQL and NoSQL databases such as MongoDB, as well as visualization tools like Redash and Superset.

MODEL CREATION

In my career, I have led complete processes of modeling and implementing machine learning models. I managed projects using the CRISP-DM methodology, covering from understanding requirements to implementation and validation in production. With this, I worked on diverse projects that span various paradigms of Machine Learning, some notable projects would be: Recommendation Systems, NLP, Reinforcement Learning with Human Feedback, Preventive Maintenance Models, Profiling Models, and prediction of human behavior. For these projects, I used a varied set of tools such as: Python, Pandas, Pytorch, Keras, Tensorflow, BigQuery, SQL, CloudSQL, MongoDB, SkLearn, Google Cloud Platform, Flask, Swagger, Looker Studio, LightGBM, XGBoost, among others.

CREATION OF CHERRYPICK LIBRARY

I developed the Cherrypick library, an innovative tool for feature selection, during my work at GyraMais. This library introduces two main approaches: 'competitive scoring', which treats feature selection as a competition among variables, evaluating the performance of models with different classifiers or metrics; and 'cherry_score', a method to evaluate variables based on their efficiency in distinguishing rows with high or low accuracy rates, similar to the distribution of ENEM scores. These approaches allowed for a more effective and reliable selection of relevant variables.

OTHER PERSONAL PROJECTS

On my GitHub, various personal projects that I have developed or are in progress are available. I highlight the industrial monitoring project, focused on predicting machine failures for preventive maintenance actions. It also includes a study on employee attrition, analyzing factors influencing the departure of employees from a company, aiming for insights beyond the conventional. Another notable project is the creation of artificial intelligence capable of composing music.

Language Skills

ADVANCED: English

INTERMEDIARY: Spanish

FIRST LANGUAGE: Portuguese