Artur Wanderley

Statistical Analysis | Machine Learning | Generative AI

@ wanderley.artur@gmail.com   🔗 Linkedin   📍 Recife, Brazil

SUMMARY

I hold a Ph.D. in Biology and possess over 15 years of experience analyzing complex data across diverse domains, including biostatistics, customer churn risk, and internet data in the context of cybersecurity. In academia, I have authored multiple scientific publications in leading peer-reviewed journals. In industry, I specialize in analyzing large unstructured datasets to solve complex business problems, developing key metrics to track business goals, and creating end-to-end machine learning applications.

EXPERIENCE

Data Scientist

Axur

📅 2020 - Present

Performing complex data analyses and developing end-to-end machine learning systems at Axur, an AI-driven cybersecurity company.

- Collaborated with cross-functional teams to conduct complex data analyses for product development and churn risk assessment.
- Developed KPIs and dashboards to monitor business performance and inform stakeholder decisions.
- Created a RAG based chatbot to assist sales teams in retrieving tailored information for prospects.
- Developed a website content classifier using embedding models, increasing the bypass rate of anti-bot mechanisms during web scraping by 47%.
- Leveraged LLMs to segment deceptive terms in suspicious URLs, enhancing the detection of fraudulent internet content by 19%.
- Applied tree-based machine learning algorithms to improve website phishing detection rates by 32%.
- Utilized Amazon EMR and Apache Airflow to automate ETL data pipelines.

Post Doctoral Researcher

Federal University of Pernambuco

📅 2018 - 2020   📍 Recife, Brazil

- Conducted exploratory data analysis (EDA) and visualization utilizing Python and R.
- Executed hypothesis testing.
- Developed predictive models forecasting plant distribution in response to climate change.
- Provided statistical support across multiple research projects.

Visiting Graduate Researcher

UCLA

📅 04/2014 - 10/2014   📍 Los Angeles, United States

Collaboration with Professor Victoria Sork during my PhD research:

- Developed predictive models to assess the impact of climate change on the genetic diversity of a threatened plant species.
- Participated in weekly seminars focused on analyzing genetic and environmental data.

EDUCATION

Specialization in Data Analysis and Data Engineering

CESAR School

📅 2019 - 2021   📍 Recife, Brazil

PhD, Plant Ecology

Federal University of Pernambuco

📅 2011 - 2015   📍 Recife, Brazil

STRENGTHS

Analytical Thinking

I use a data-centric approach to validate new ideas and challenge assumptions based solely on common sense.

Problem-Solving Skills

My approach focuses on solving real-world problems with well-suited tools, rather than implementing technologies solely because they are trending.

Collaboration and Teamwork

I am recognized for being collaborative and approachable, consistently contributing to team projects and fostering a supportive work environment.

SKILLS

Statistical analysis   Machine learning

Generative AI   RAG   Vector database

Airflow   AWS   Python   SQL

SCIENTIFIC PUBLICATIONS

Full publications
Link to Google Drive here.

INTERESTS

Data Science

My passion lies in analyzing data to uncover insights and implement practical solutions.

Surfing

Surfing enthusiast in the warm waters of Northeast Brazil.