

WILLIAM WAYN

Lead Data Scientist | Machine Learning Expert

 blog.wayn.com.br

 Brazil

 willzwayn@gmail.com

 github.com/WillzWayn

 linkedin.com/in/willzwayn

PROFILE

Experienced Lead Data Scientist with over 5 years in machine learning and data analytics, currently driving innovative solutions at BTG Pactual. Expertise in deploying production-grade models and optimizing data-driven marketing strategies. Proven ability to manage cross-functional teams and deliver impactful projects across finance and tech industries.

SKILLS

Experience Years:

Python 6 Years
SQL 5 Years
AWS 5 Years
Linux / Bash 5 Years

Tech List:

Data Science and ML:

Python, PySpark (Hadoop), Mlflow, Hugging Face, LLaIndex, CrewAI, LangChain, Generative AI, Airflow, CI/CD, Docker, MLOps.

Cloud Services:

AWS, Amazon Bedrock, Sagemaker, ECS, ECR, Lambda, Athena, Glue, GCP

Others:

Django, FastAPI, PostgreSQL, GIS, GDAL, RASTERIO, QGIS

AWARDS

AWS

AWS Community Builder

[AWS community-builders-directory](#)

Recognized by AWS for active contribution and leadership in the community. Engaged in collaborative initiatives, knowledge sharing, and use of AWS technologies for community impact.

BCG

TOP #2 in Boston Consulting Group - Gamma Challenge - 2021

[Linkedin Post](#)

During the BCG hackathon, I addressed Brazil's domestic violence problem. I collected additional data on denunciations, aggressions, and homicides, and analyzed 5,568 Brazilian municipalities. I developed a model to identify anomalies between denunciations and aggressions, enabling targeted interventions. The solution helped the Igarapé Institute make a meaningful social impact.

EDUCATION

2014-2019

Bachelor in Physics

[UFRRJ - Brazil](#)

I received a scholarship to teach reinforcement classes in physics and calculus. Additionally, I was experimental Physics Monitor, where I had my first exposure to the field of machine learning. I assisted students with theory behind graphic construction, curve-fitting methods, including regression

2020

Bootcamp - Data Science & Machine Learning for Business

[Tera](#)

COURSES

2023

On-site Course in Master's Program in Data Science

[UNIRIO - Brazil](#)

Completed Deep Learning course as external student. Presented "Transformer is all you need" article on transformers architecture.

2021

Introduction to Machine Learning

[Coursera - Duke University](#)

Completed a comprehensive Machine Learning course, consisting of 6 modules, with a total duration of 21 hours. The course covered a wide range of topics, including: Introduction to Machine Learning, including logistic regression, multilayer perceptron, and deep learning; Mathematical basis of learning deep networks, including optimization techniques such as gradient descent and stochastic gradient descent; Convolutional Neural Networks (**CNNs**) and their applications, including image analysis; Natural Language Processing (NLP) using neural networks, including word embeddings, recurrent neural networks, and long short-term memory (**LSTM**) models; **Transformer Networks for NLP**; **Reinforcement Learning**, including Q Learning and **Deep Q Learning**

2020

Google Cloud Platform Big Data and ML Fundamentals

[Coursera](#)

Implemented big data and ML solutions on GCP, covering data-to-AI lifecycle. Created streaming pipelines with Dataflow/Pub/Sub, explored ML model options, and built pipelines with AutoML. Proficient in leveraging GCP for data-driven insights.

- Coursera - Object Detection with Amazon Sagemaker
- Coursera - Using TensorFlow with Amazon Sagemaker
- Datacamp - Intermediate Python
- Datacamp - Introduction to TensorFlow in Python

2019

IBM-Python for Data Science
Certification link

CERTIFICATIONS

AWS

AWS Certified Machine Learning - Specialty

Credential Verification

Certified in AWS Machine Learning (ML) services, competent in building, training, tuning and deploying ML models in the AWS cloud. Skilled in utilizing both pretrained models and custom models from open-source frameworks to derive insights from AWS ML services.

Airflow

Astronomer Certification for Apache Airflow Fundamentals

Credential Verification

Certified in Apache Airflow Fundamentals by Astronomer, proficient in Airflow architecture, task life cycle, and scheduling processes. Skilled in recommending architecture needs, settings, and design choices for data pipelines. Capable of task triggering, debugging, and monitoring through UI views.

EXPERIENCE

- 11/2020 – Present **BTG Pactual - Investment Bank** **Data Scientist - ML Engineer**
- **Lead Data Scientist - Fraud / AML Prevention Team (01/2023 – Present)**
 - Managed a team of three professionals in the Anti-Financial Crime team, ensuring successful project execution and team collaboration.
 - Designed a project to prevent credit loops and self-financing, focused on reporting to Anti-Money Laundering (AML) efforts based on graphs.
 - Successfully implemented a novel model to detect and prevent "brand fraud", a pervasive type of fraud where fictitious companies disguise themselves as legitimate organizations. This innovative solution, built on a fundamental model for business entities, virtually eliminated the presence of fraudulent companies, achieving a staggering 100% reduction in brand fraud cases, thereby saving the organization from substantial financial losses and reputational damage.
 - **Machine Learning - MLOps Team (06/2022 – 01/2023)**
 - Designed and developed a machine learning pipeline from scratch, designing a GitOps-based pattern for code management, CI/CD, and automated deployment. This end-to-end pipeline enabled seamless experimentation and retraining, leveraging MLflow as a model registry, Feast as a feature store, and Great-Expectations for data quality control. Additionally, implemented automatic monitoring of model and business metrics in Grafana.
 - **Data Scientist - Propensity Score Team (11/2020 – 06/2022)**
 - Developed and deployed a recommendation system model based on customer consumption and navigation history, resulting in a 31% increase in conversion rates in marketing campaigns.
Python / PySpark (Hadoop) / Docker / Deep Learning / NLP / Hugging Face / SQL / Mlflow / Airflow / CI/CD / MLOps / AWS / ECS / ECR / Lambda / Athena / Glue / Sagemaker / Fast Api
- 2022/06 – 2024/01 **Expert | Teacher** **Tera**
- Taught data science classes, including "Complex Business Problem Solving" and "Deploying Machine Learning Models in Production".
Problem Solving
- 2020.06 – 2020.12 **Freelancer** **Data Scientist Consultant**
Own Company
- Worked as a consultant for Santander, a large bank in South America. Deployed attributions models based on Markov chains and Shapley value using Google Analytics raw data.
 - Developed strategies to prevent customer churn, enhancing customer retention efforts.
Python Docker / Marketing Mix Modeling (MMM) / Google Cloud Platform (GCP)
- 2019.12 – 2020.12 **SciCrop** **Data Scientist**
- As part of a consultancy for a major international Sugar Cane company, I led the end-to-end implementation of a serverless AWS project.
 - Developed a solution for detecting diseases in sugarcane fields using geolocated images from drones.
 - Key responsibilities and achievements included:
 - Interfaced with stakeholders to understand business requirements.
 - Established a methodology for image labeling in collaboration with the operational team.
 - Built a proof-of-concept (POC) deep learning model with the Django API.
 - Iteratively refined labels based on model feedback.
 - Implemented CI/CD automation pipeline in GitLab for model training and versioning using MLFLOW.
- Python / MLOPS / AWS / GIS / GDAL / RASTERIO / QGIS / Python / PostgreSQL / SQL / Django / Deep Learning / Docker / Git
- 2019.01 – 2019.12 **Brazilian NGO - Instituto Casa Viva** **Physics and Math Teacher**
- Taught Physics and Math to low-income students in a social pre-university program at a local church, preparing them for university entrance exams.
 - Developed and delivered targeted lessons, resulting in improved academic performance and encouraging students to pursue higher education.

LANGUAGES

- Portuguese **Native**
Fluent in speaking, reading, and writing.
- English **Advanced**
Proficient in professional and technical communication.
- Spanish **Intermediate**
Competent in conversational and basic technical contexts.