



Bernardo Galvão

📍 Madeira, Portugal

MLOps Engineer

bgalvao

bgalvao

bcbgalvao

Data Scientist turning into a full-fledged MLOps Engineer. My experience in Data Science and Software Engineering is resulting in the perfect positioning for this new area of expertise. Currently acquiring DevOps skills and counting. Hire me if you need someone to operationalize the Machine Learning Operations in your company and let the Data Scientists focus on their modelling task: I'll handle the operations.

Skills:

MLOps

CI/CD

Python

Kubernetes

Scikit-Learn

Grafana

AWS

Experience

Cardo AI cardoai.com

Feb 2023 to Present

MLOps Engineer Remote

Machine Learning Engineer with AWS and Kubernetes

- Created a Kubernetes CronJob connected to a Slack bot that switches off KubeFlow notebooks, therefore cutting AWS costs by releasing notebook-bound EC2 resources.
- Developed Argo Workflow for monitoring of models in production, using NannyML, Kubernetes, AWS SQS and S3 Event notifications, allowing for observability of performance and feature drift.
- Enhanced data processing workflows in AirFlow and added unit tests to AirFlow DAGs.
- Deployed and managed Grafana, Prometheus and Loki for general monitoring of K8s resources.
- Added model pre-deployment tests in CI/CD pipelines to ensure model, training and data consistency.

MRIcons Ltd. mricons.eu

Oct 2021 to Present

MLOps Engineer (Contractor) Remote

- Developed Docker Swarm stacks for deployment of MLOps services, including a model registry (MLFlow); a reverse-proxy (Traefik) with an authentication service; a workflow orchestrator (Prefect).
- Designed an on-prem MLOps platform to provide end-to-end services for frictionless training, validation, deployment and monitoring of ML models based on Open-Source software, including a CI/CD component for ML.
- Designed a workflow template using DVC for Data Scientists to confidently develop datasets and models with unit testing integrating a CI/CD pipeline written in GitLab CI/CD specification.
- Wrote tests and GitLab CI pipelines for deployment of thoroughly tested models on their data and performance. Wrote Gitlab CI pipelines for automated deployment of Docker Stacks.
- Optimized performance of data queries and transformations using optimized memory formats Apache Arrow.

Champalimaud Foundation [Computational Clinical Imaging Group]

Jan 2020 to Jul 2021

Research Fellow Lisbon, Portugal

- Deployed MLFlow server for experiment tracking and experiment data collection.
- Conducted research on performance feature selection methods applied to radiomics for clinical decision-making.
- Implemented Target Shuffling robustness check and Nested Cross-Validation procedure for use by other members in the team.

Jun 2018 to Feb 2020

Madeira Interactive Technologies Institute

Data Scientist

- Implemented an object detection viewer and trained a marine-species object detector via transfer learning.
- Automated dataset processing, model training and Tensorflow model format conversion to TensorFlowJS and TensorFlow Lite using Docker, Python and Bash.
- Provided a user-friendly CLI for pulling, configuring and training models from TensorFlow's Object Detection API.
- Implemented process for cooperative collection, annotation and augmentation of training images.
- Analyzed social media data activity levels per location, including profiling points of interest on Madeira island via topic modelling of TripAdvisor reviews with Latent Dirichlet Allocation.
- Performed data wrangling and analysis to assess performance of low-cost air quality sensors.

Eyeware eyeware.tech

Nov 2017 to Feb 2018

Software Engineering Intern Lisbon, Portugal

- Developed a raycasting prototype in order to produce an attention heatmap on a 3D object using NumPy and VTK in Python.

Education

Nova Information Management School (IMS) novaims.unl.pt

Jan 2015 to 2017

MSc, Data Science and Advanced Analytics

Católica-Lisbon School of Business and Economics

2011 to 2014

BSc, Economics

Projects

nodevogithub.com/bgalvao/nodevo

An implementation of Genetic Programming in Rust.

Rust

a-priorigithub.com/bgalvao/a-priori

Extracting association rules from a market-basket dataset using the A-Priori counting strategy. This is a university project coded in Java for Hadoop's MapReduce.

Java

Publications

Prediction of Prostate Cancer Disease Aggressiveness Using Bi-Parametric MRI Radiomics

Cancers 2021

pubmed.ncbi.nlm.nih.gov/34885175/**A Parallel and Distributed Semantic Genetic Programming System**

2017 IEEE Congress on Evolutionary Computation (CEC) 2017

ieeexplore.ieee.org/document/7969304

Languages

Portuguese Fluent

English Fluent