Bruno Jacob

Goleta, CA | (805) 617-6974 | bruno.mflab@gmail.com | github.com/brunopjacob | US Permanent Resident

Technical Skills

PROGRAMMING LANGUAGES: Python | Fortran | Julia | SQL | MATLAB | C++ | Bash | Scala

DATA SCIENCE: NumPy | Scikit-Learn | Imaging Processing | Data Visualization (Matplotlib) | Pandas | Spark

MACHINE LEARNING: Spark MLLib/ML | TensorFlow | Keras | NLP | Deep Learning

LANGUAGES: English, Portuguese (native speaker)

Education

PhD in Computational Science and Engineering | University of California, Santa Barbara | Sep 2021

MSc in Mechanical Engineering | Universidade Federal de Uberlandia, Brazil | Aug 2015

BSc in Mechanical Engineering | University of Illinois at Urbana-Champaign & Universidade Federal de Uberlandia, Brazil | Aug 2014

Experience

CARPE DATA **Mar 2021 – Present**
Machine Learning Engineer

- Developed and maintained a Scala/Spark data ingestion pipeline using best data practices (Git, Jenkins, CI/CD).
- Analyzed large datasets (100s of TBs) to identify, parse and extract structured and unstructured data using Spark.
- Deployed and prototyped big data pipeline using Elastic MapReduce (EMR), S3, Terraform and AWS Lambda.
- Developed tools to resolve entity resolution and machine learning problems using Spark ML and GraphX.

UNIVERSITY OF CALIFORNIA, SANTA BARBARA **Sep 2017 – Mar 2021**
Graduate Research Assistant (Advisor: Prof. Linda Petzold)

- Built a parallel (MPI, OpenMP, Cuda) C++ software to investigate the effects of stochastic dynamics in biophysical systems subject to multiple chemical reactions and fluid flows using coarse-grained physics (DPD, SPH) and finite element/ finite differences/ spectral methods.
- Developed supervised machine learning models (logistic regression, SVM, decision trees, and random forests) to optimize chemical reactions and accelerate the fabrication of quantum materials by over 90%.
- Designed natural language processing (NLP) models to extract relevant data from 1K+ of scientific journals with over 92% accuracy, replacing previous manual literature searches and saving 100s of hours of tedious manual searching.

ARGONNE NATIONAL LABORATORY **May 2018 – Sep 2018**
Associate Visiting Researcher

- Developed software in Julia/Python that predicts power line overload, desynchronization and blackouts and performed data analysis on 10's of GBs of data using Julia and SQL.
- Effectively presented results to a non-scientific audience of stakeholders to secure research funding.

RELEVANT COURSEWORK (Coursera/Stanford/DataCamp)
Machine Learning Specialization | Neural Networks & Deep Learning | Quantitative Modeling Fundamentals | Natural Language Processing with Classification and Vector Spaces | Structuring Machine Learning Projects | OOP in Python

Projects

STOCHSS (STOCHASTIC SIMULATION SERVICE) github.com/StochSS/stochss
Contributor in building a web-based service for modeling, simulation, and analysis of a wide range of mathematical, biological, and biochemical systems; written with C and C++ and developed the UI with Python.