Marcos de Souza Oliveira

Recife, Brazil marcosd3souza@gmail.com +55 081 98784-4363 marcosd3souza.github.io
marcos-d-souza marcosd3souza

Summary

I am a highly skilled researcher and professional with over 14 years of experience in software development and Artificial Intelligence (AI), Machine Learning (ML), and data analytics. I have led and contributed to a wide range of projects, from agile development teams to large-scale international collaborations, holding roles such as Tech Lead, AI/ML specialist, Data Scientist, and Software Engineer. I am particularly drawn to unsupervised learning and exploratory data analysis.

Interests

My main interest is to develop meaningful, cost-effective and explainable AI models through unsupervised deep learning, exploratory data analysis, and optimization, specifically deep clustering, deep graph clustering, evolutionary algorithms and dimensionality reduction methods.

Education

PhD **Federal University of Pernambuco - UFPE**, Machine Learning — 2020 – 2024
Title: Unsupervised Feature Selection and Deep Subspace Clustering for Exploratory High-Dimensional Cluster Analysis
URL: https://repositorio.ufpe.br/handle/123456789/62451

MSc **Federal University of Pernambuco - UFPE**, Machine Learning — 2016 – 2018
Title: Unsupervised Feature Selection Methodology for Clustering in High Dimensionality Datasets
URL: https://repositorio.ufpe.br/handle/123456789/33642

Spec **University of Pernambuco**, Software Engineering — 2015 – 2016
Title: Proposta de Arquitetura Utilizando o Paradigma SOA para o Avatar Educação
DOI: 10.25286/repa.v3i1.528

BS **Faculdade dos Guararapes**, Computer Science — 2011 – 2014

Teaching Experience

CESAR School, Specialization, Professor — Recife, Pernambuco, Brazil, 2022 - 2023

- machine learning
- unsupervised learning
- dimensionality reduction

URL: https://www.cesar.school/especializacao/analise-de-dados-e-ia

Universidade Tiradentes - UNIT, Graduation, Professor — Recife, Pernambuco, Brazil, 2020

- algorithms and compiler components
- data structures
- Problem-Based Learning methodology

URL: https://www.unit.br/analise-e-desenvolvimento-de-sistemas

Relevant Experience

Inatel, AI/ML Specialist — Minas Gerais, Brazil (remote)

- Development of statistical and AI/ML models applied to industrial process

- Prospect new AI/ML projects by writing pre-proposal/pre-sale and work plan documents
- Conduct studies and analysis of AI/ML frameworks, data, platforms and tools
- Publish results in top conference/journals (Scopus indexing)

2023 - current

CESAR, Data Scientist — Recife, Pernambuco, Brazil — 2019 - 2023

- Development of explainable classifier (tree-based) for predicting employee turnover
- Development of anomaly detection models to identify outliers/crashes in applications that observe data consumption over time (time series)
- Design and implementation of the queue system simulator to perform data ingestion for AI/ML models
- Development of route suggestion system based on clustering for Samsung technicians (optimization of traveled distance to meet all demanded attendances)
- Development of optimization models to perform the best organization of queues in Samsung's assistance by evolutionary algorithms (genetic / PSO)

CESAR, Software Engineer — Recife, Pernambuco, Brazil — 2018

- Implementation of NLP classifier to detect furious feedback in the consumers text messages (Samsung concierge)

Stefanini, Software Engineer — Recife, Pernambuco, Brazil — 2017

- Developing mobile applications for pattern recognition engines (image document segmentation).

Informatics Center, UFPE, Machine Learning Scholarship — Recife, Pernambuco, Brazil — 2016

- Dedicated to research work for the master's degree in Machine Learning
- Unsupervised Feature Selection Framework

Selected Publications

Deep Contrastive Variational Subspace Clustering — 2025

Oliveira, M. D. S., Queiroz, S. R. M., Zanchettin, Cleber., Carvalho, F. A. T.

Neurocomputing

10.1016/j.neucom.2025.130901

FastAiAlloc: A real-time multi-resources allocation framework proposal based on predictive model and multiple optimization strategies — 2023

Oliveira, M. D. S., et al.

Future Generation Computer Systems

10.1016/j.future.2023.08.014

Unsupervised feature selection method based on iterative similarity graph factorization and clustering by modularity — 2022

Oliveira, M. D. S., Queiroz, S. R. M., Carvalho, F. A. T.

Expert Systems with Applications

10.1016/j.eswa.2022.118092

Unsupervised feature selection methodology for clustering in high dimensionality datasets — 2020

Oliveira, M. D. S., Queiroz, S. R. M.

Revista de Informática Teórica e Aplicada

10.22456/2175-2745.96081

Technologies

Languages: Python, Matlab, Java, SQL, JavaScript

Technologies: Scikit-learn, Pytorch, Tensorflow, NetworkX