Vitor N. Cabral de Oliveira

vitorncabral@gmail.com | +55 81 99847 - 4134 | github.com/vitornegromonte
Data Scientist
linkedin.com/in/vitornegromonte

EDUCATION

Federal University of Pernambuco August 2022 - August 2026
*Bachelor of Science in Statistics * *GPA 6/10*

WORK EXPERIENCE

Geraia Recife, Pernambuco, Brazil
Artificial Intelligence Engineer/Researcher November 2023 - Present

- Research Group on Generative AI Applications;
- Conducting research in Generative AI, focusing on evaluating Large Language Models (LLMs) in Portuguese;
- Working on improving Retrieval-Augmented Generation (RAG) through the application of advanced mathematical methods and probabilistic techniques, aiming to enhance model performance and reduce hallucinations in critical systems.
- Key-words: *Python, Pytorch, Natural Language Processing, Optimization, Deep Learning, RAG, LLM.*

Redduo.ai Recife, Pernambuco, Brazil
Data Scientist – part time November 2023 - July 2024

- Worked as a Data Scientist, conducting data analysis to drive business intelligence initiatives and developing automations using AWS services (Lambda, Cognito, etc.). Contributed to the development of core AI models for business applications, focusing on model optimization and performance enhancement.
- Key-words: *Python, Pytorch, Machine Learning, Business Intelligence, Predictive Analysis, Data Engineering*.

INES – National Institute of Software Engineering Recife, Pernambuco, Brazil
Junior Researcher July 2023 - May 2024

- Researching in quantitative methodologies for enhancing web and mobile accessibility for user in the Autistic Spectrum Disorder (ASD) with a focus in urban mobility applications.
- Key-words: *Data Analysis, Human-Computer Interaction, Software Engineering.*

PROJECTS

EnergIA.py August 2024 – Present
Open-source Python package for GPU energy optimization during Artificial Intelligence Training

- An open-source Python package designed for GPU energy optimization during artificial intelligence training. EnergIA.py aims to reduce the energy consumption and carbon footprint of AI models by optimizing GPU usage, making AI training more efficient and environmentally friendly. Utilizes classical machine learning and optimization techniques to achieve energy-efficient GPU utilization.

Detecting Parkinson's Disease Using Computer Vision Techniques August 2024 – September 2024
Campus Pary Nordeste – 2024 keynote

- Designed and developed a project leveraging computer vision techniques to facilitate fast and cost-effective medical diagnosis. Utilized OpenCV for image preprocessing, and PyTorch, TensorFlow, and CUDA for model training, while optimizing hyperparameters with Optuna. The project achieved an average accuracy of 90% across multiple models, with the best model reaching 95% accuracy, significantly outperforming the previous best results of 85%.
- Key-words*: OpenCV, Computer Vision, Image Processing, Pytorch, Tensorflow, Nvidia CUDA*

Breast Cancer Detection March 2024 – April 2024
A CNN Ensemble Methodology for Enhancing Breast Cancer Diagnosis

- Developed a project to improve the accuracy and reliability of breast cancer diagnosis using a Convolutional Neural Network (CNN) ensemble methodology. Utilized an ensemble of CNN models to analyze medical images, leveraging the strengths of multiple models to enhance diagnostic performance. The result had accuracy of approximately 92%.
- Key-words*: Python, Pytorch, Computer Vision, Classification, Image Processing*

SKILLS

PROGRAMMING LANGUAGES
3 years: Python, R, SQL
3-6 months: C++, Go, Mojo

LANGUAGES
English: Proficient or Fluent
Spanish and French: Basic

TECHNOLOGIES
Python, C++, R, SQL, PostgreSQL, PowerBI, Google DataStudio, Pandas, NumPy, SciPy, scikit-learn, PyTorch, TensorFlow, Pytorch Lightning, Keras, Optuna, CUDA, OpenCV, Cloud AWS