

Yanka Ribeiro

📍 Maceio, Brazil ✉ ynk9a@outlook.com ☎ +55 82 9 9991 5609 🔗 <https://www.linkedin.com/in/yanka-ribeiro/> 📁 <https://github.com/yrribeiro>

EXPERIENCE

Data Scientist Intern | IBM Brazil | Sao Paulo, BR | April 2022 – December 2022

- Built a Real-time opportunity evaluation engine to identify the best parameters to be observed by a bank employee for cost data labelling, aiming to get a better understanding of the bank's expenses and portfolio.
- Developed and deployed multiple propensity models by using machine learning and deep learning methodologies including Random forest, Multi-layer perceptron and SVM, to determine the propensity of an expense to be part of a determinate group, and applied the models to the existing database to generate hints for advisors and help the portfolio team create personalized assets show-offs.
- Built a combined version of an NLP and supervised ML model to determine the most common budget category associated with each expense group, with a precision of 92%.

Data Scientist – NLP Specialist | Centro de Inovação EDGE | Maceio, BR | January 2021 – February 2022

- Developed a rule-based matching engine using spaCy, to create a recommendation system.
- Performed data augmentation to generate data about medical equipment specification, needed for the training step of the recommendation system.
- Architected a fine-tuned BERTimbau (Portuguese version) for matching and extracting more complex and dynamic features to compare with an existing database.

EDUCATION

BSc. Computer Science | Minor in Artificial Intelligence | Federal University of Alagoas | Maceió, BR | 2023 | 3.34

INVOLVEMENT

Member | Federal University of Alagoas | Hacker Academy Outreach Program (AcHA) | May 2022 – Present

- Write articles to be released in a biweekly e-journal, about interesting topics inside Cybersecurity and Computation in general.

Secretary | Federal University of Alagoas | Student Academic Office | July 2021 – July 2022

- Implemented an organization system of filing and tracking for files management on cloud.
- Organized office tasks and inventory, maintained records of all meetings and students demands.
- Provide reception to freshman students, giving advices, best practices and campus tours.

Student Volunteer | Federal University of Alagoas | IEEE Computer Society | September 2022 – Present

Team Member | Federal University of Alagoas | GEMA – Artificial Intelligence Team | March 2023 – Present

- Member of the first A.I specialized team inside the outreach program (GEMA) that prepares computing students for programming marathons.
- Our team represents the university in online AI and data challenge competitions

CERTIFICATIONS

AI Associate Data Scientist | IBM | 2022

This badge earner demonstrates an understanding of artificial intelligence (AI) concepts, data science fundamentals, and the IBM strategy for AI. This includes knowledge about the Watson products and services, and data visualization.

Machine Learning Specialist | IBM | 2022

The earner can install and use tools like Python, Jupyter Notebooks, SQL, and Docker, as well as connecting to various open source data sources. Understands the requirements, terms and concepts associated with successfully developing and deploying AI solutions in their enterprises

Azure Fundamentals | DIO | 2023

Introduction to AWS Cloud Computing | DIO | 2023

Developing Serverless Applications with AWS | DIO | 2023

Data Engineering with Azure and AWS | DIO | 2023

Data Science Methodologies | IBM | 2022

Understands the essential steps used in data science business and research problem solving: problem definition, collecting and analyzing data, building relevant models and understanding model deployment results

Agile Explorer | IBM | 2022

Foundational understanding of Agile values, principles, and practices that allows them to initiate an Agile conversation with co-workers and colleagues

First Place at Programming Marathon | Federal University of Alagoas | 2019

First place in the Programming Marathon during my first year as a computer science student, while competing with other students, including veterans. The main programming language used was C, and the questions tested concepts of data structures and programming logic.

SKILLS

Machine Learning: Python (scikit-learn, pandas, NumPy), TensorFlow, Pytorch

Data Processing: SQL

Data Visualization Tools: R Studio, Latex, Excel

Cloud Platforms/Environments: AWS, GitHub

Database: MySQL

Programming languages: Python, C/C++, Java, R, JavaScript

SUMMARY

As an aspiring data scientist/machine learning engineer, I am dedicated to applying my analytical skills to derive insights from complex data sets, with a passion for problem-solving and a willingness to learn and grow in the field.