MACEIO, BRAZIL
yrribeiro.github.io
ynk9a@outlook.com

MACHINE LEARNING AND COMPUTER VISION ENGINEER

Yanka Ribeiro

TECH STACK

python, c++, pandas, numpy, scikit-learn, pytorch, opencv, seaborn, matplotlib, pyspark, google cloud (big query, vertex ai), mysql, git, docker, flask.

RESEARCH EXPERIENCE

Google Explore CSR LATAM Program, Remote, Advisor Phd. Enzo Ferrante (CONICET Argentina) — *Leveraging Vision Language Models as metadata generators for medical imaging datasets*

february, 2024 - july, 2024 **(in progress)**

- leverage the open-vocabulary capacity of current Vision Language Models (VLMs) and evaluate their potential to describe and annotate artifacts and other types of contextual metadata in X-ray medical imaging datasets, to enable future analysis of biases in computer aided diagnosis, which take into account artifacts as potential reasons.

National Council for Scientific and Technological Development (CNPq), Remote (UNCISAL), Advisor Phd. Raquel Cabral (UFAL Arapiraca) — *Uma ferramenta computacional baseada em machine learning para registro e análise de potenciais evocados auditivos nos domínios do tempo e da frequência*

january, 2024 - december, 2026 **(in progress)**

- develop a computational tool capable of recording and evaluating auditory evoked potentials in the time and frequency domains, thus assisting professionals in generating more precise reports for exams. This aims to facilitate the early diagnosis of auditory alterations, including the possibility of a more refined evaluation of central auditory processing, with extensive options for electrophysiological investigation. Certain conditions, such as deafness, could be explored from the child's birth.

Autonomous (Bachelor's Thesis), Hybrid (UFAL), Advisor Phd. Fabiane Queiroz (UFAL Maceió) — *Evaluation of Deep Metric Learning Methods for the Diagnosis of Human Visceral Leishmaniasis*

september, 2023 - december, 2023 (three months)

- build and evaluate the impact and effectiveness of deep metric learning methods in accurately diagnosing human visceral leishmaniasis using microscopic images by: accentuating areas of relevance within the images and segmenting these images into smaller patches, comparing 4 pre-selected deep metric learning algorithms to pinpoint the most

effective models for extracting features and configuring a supervised classification algorithm to categorize images based on the data extracted from the metric learning models.

computer vision, image processing, image classification, python, pytorch, cnn, svm, pca, deep metric learning

São Paulo Research Foundation (FAPESP), Remote (UNIFESP), Advisor Phd. Rubens Belfort Jr. — *Uso de algoritmos de Deep Learning para rastreio de Toxoplasmose ocular em imagens de retina*

january, 2023 - march, 2023 (three months)

- develop and validate a computer vision algorithm for tracking ocular toxoplasmosis in the Brazilian population using the BR-OPHTSET database, a collection of colored retina photos featuring Brazilian patients and patients from the Uveitis Outpatient Clinic at the Federal University of São Paulo (UNIFESP).
- the most commonly used neural networks in computer vision, such as ResNet, VGG, DenseNet, and EfficientNet, were applied and compared with different optimizers.

computer vision, image segmentation, python, pytorch, cnn, residual nets

UFAL Office for Research and Postgraduate Studies (PROPEP), Remote (UFAL/BrAIN), Advisor Phd. Aydano Machado — *O uso da Inteligência Artificial para qualificar o resultado do cálculo do Poder da Lente Intraocular em portadores de catarata*

september, 2021 - august, 2022 (twelve months)

- creation of decision support systems based on computational models that qualify the result of classical formula for calculating the power of the intraocular lens (SRK/T) implanted in phacoemulsification surgery in patients with senile cataracts, using information obtained from the IOL Master 500® optical biometer and the results of classical formulas for calculating the degree of the intraocular lens.
- constructed and evaluated predictive models based on decision trees and Bayesian networks for the mentioned formula.
- developed a mechanism to detail the inference performed by each model after prediction, making it understandable for the surgeon.

supervised algorithms, random forests, bayesian nets, tabular data wrangling, python

WORK EXPERIENCE

IBM, Remote (Sao Paulo HQ) — *Data Scientist Intern*

april, 2022 - december, 2022 (eight months)

- full stack experience in machine learning: built an end to end production evaluation engine to optimize cost data labeling for bank employees, enhancing insights into the bank's expenses and team portfolio.
- achieved 90% precision, on average, with a variety of supervised and unsupervised ML models to determine the most common budget category associated with each expense pattern.

Edge AI Laboratory, Maceio — *Machine Learning Engineer, NLP Team*

january, 2021 - february, 2022 (one year)

- designed and implemented a rule-based NLP engine for automatic extraction of key specifications from lengthy tender documents, later used in a recommendation system for a company to optimize their bid process in public tenders.
- responsible for fine-tuning BERTimbau (PT-BR) for matching and extracting more complex and dynamic features related to CT equipment technical specifications.
- successfully reduced the chances of bid losses due to inaccurate equipment recommendations by 65%, leading to improved win rates in competitive tenders.

EDUCATION

Federal University of Alagoas, Maceio — *BSc. Computer Science*

september, 2019 - december, 2023

- teaching assistant, twice, for the Computer Science department.
- member of The Student Academic Office of Computer Science and Engineering (DIACOM).
- volunteer in outreach programs: AcHA (cybersecurity), IEEE (research and robotics) and GEMA (competitive programming).
- two undergraduate research internships in the field of biomedical computing, focusing on ml/dl applications in ophthalmology.
- scholarship holder at the Office of International Affairs.
- 8.5 / 10 gpa