

# RODRIGO YEPEZ-LOPEZ

Bethesda, MD | rodrigoyepezm@gmail.com | [Github](#) | (240) 515-0030 [LinkedIn](#)

I am a Machine Learning Engineer and Data Scientist with hands-on experience applying deep learning to healthcare, biomedical, and scientific datasets. I've built multimodal AI systems, large-scale SQL and PySpark pipelines, and HPC-trained models for dementia analytics, pulmonary forecasting, AI cancer classification and proton radiotherapy image enhancement. I am proficient in Python, NumPy, Pandas, scikit-learn, and PyTorch, with extensive experience developing end-to-end analytics and ML workflows. I have four peer-reviewed publications in machine learning I bring strong foundations in GenAI, LLMs, reproducible data pipelines, statistical modeling, feature engineering, model evaluation, and communicating complex technical insights to cross-functional teams.

## EDUCATION

### University of Maryland, Baltimore County (UMBC)

MS in Computer Science (3.6/4.0)

Baltimore, MD

May 2025

### American University

BS in Computer Science

Washington, DC

May 2023

## SKILLS

**Programming:** Python, SQL, C++, Pandas, Git, Docker, REST APIs, Scikit-Learn, PyTorch, Keras, PySpark, AWS

**Machine Learning & Modeling:** Deep Learning, CNN, Transformers, Object Detection, Segmentation, SfM, Depth Estimation, NeRFs,

**Soft Skills:** Agile Development, Technical Documentation, Cross-Functional Teamwork, Fluent (English, Spanish, French)

## WORK EXPERIENCE

### UMBC, Division of Information Technology

Baltimore, MD

General Associate

May 2025 – Present

- Lead the development of real-time transcription and visualization pipelines (Linux, SQL, Python, LLMs, AWS Redshift, Tableau, Power BI) to analyze classroom conversations and support research on effective classroom interactions.
- Collaborated with cross-functional Agile teams to troubleshoot and optimize deployed AI solutions.

### UMBC, College of Engineering and Technology

Baltimore, MD

Graduate Research Assistant

August 2023 – May 2025

- Architected multimodal AI systems that integrated audio and text data through feature extraction, improving dementia classification accuracy.
- Built and maintained a 10TB SQL database to improve AI cancer classification.
- Collaborated weekly with cross-disciplinary teams to align technical solutions with research and business needs, standardizing diverse data formats and units to ensure consistent database integration.
- Documented technical processes to ensure reproducibility, accessibility, and compliance across research workflows.

### American University, Big Data Lab

Washington, D.C.

Undergraduate Research Assistant

September 2021 – May 2023

- Developed scalable pipelines using NASA imagery using Linux, Python, SQL, and Spark, improving Nasa cloud classification by 20% and enhanced weather forecasting models.
- Independently designed and implemented an LSTM model on medical signal data to predict lung recovery up to 8 hours in advance, enabling doctors to make timely ventilator removal decisions.
- Built a real-time MLOps data drift monitoring system with Python, TensorFlow, and Spark to detect shifts in streaming data and auto-trigger model retraining.

### Big Data Research Experience for Undergraduates (REU), UMBC

Washington, D.C.

Intern

June 2021 – August 2021

- Optimized compact neural networks through hyperparameter search (C, Linux, Python), reducing parameters by 30% while maintaining accuracy.
- Benchmarked model performance and created visualizations to communicate results to technical stakeholders.

## PUBLICATIONS AND SELECTED CONFERENCES

- I. Whitehouse, **R. Yepez-Lopez**, R. Corizzo (2023) "Distributed Concept Drift Detection for Efficient Model Adaptation with Big Data Streams" *2023 IEEE International Conference on Big Data (BigData)*, Los Alamitos, CA, USA

- Roberto Corizzo, **Rodrigo Yepez-Lopez**, Sebastien Gilbert, Nathalie Japkowicz. (2022) “LSTM-based Pulmonary Air Leak Forecasting for Chest Tube Management” *2022 IEEE International Conference on Big Data (Big Data)*
- Leah Ding, Roberto Corizzo, Colin Bellinger, Nancy Ching, Spencer Login, **Rodrigo Yepez-Lopez**, Jie Gong, Dong L. Wu. (2022) “Imbalanced Multi-layer Cloud Classification with Advanced Baseline Imager (ABI) and CloudSat/CALIPSO Data” *2022 IEEE International Conference on Big Data (Big Data)*
- Sokhna A. York, Alina M. Ali, David C. Lashbrooke, **Rodrigo Yepez-Lopez**, Carlos A. Barajas, Matthias K. Gobbert, Jerimy C. Polf. (2021) “Promising Hyperparameter Configurations for Deep Fully Connected Neural Networks to Improve Image Reconstruction in Proton Radio-therapy” *2021 IEEE International Conference on Big Data (Big Data)*