

Aires Augusto Miguêns

☎ 303 725 8668 | ✉ amiguens@wpi.edu | [Linkedin](#) | [Github](#) | [Website](#)

Objective

Master's student in Data Science. Experience in machine learning, deep learning, and predictive analytics. Passionate about applying AI to real-world problems, particularly in healthcare. Currently researching AI-driven healthcare solutions with Prof. Emmanuel Agu at WPI.

Education

- Worcester Polytechnic Institute (WPI), Worcester, MA | Master of Science in Data Science | Expected: May 2026

Current GPA = 3.86

- Institute of Technology of the University of Luanda, Angola | BS in Telecommunications Engineering | Graduated: 2022

Skills & Tools:

- Programming: Python, R, MATLAB, PostgreSQL, C programming language
- Machine Learning & AI: TensorFlow, PyTorch, Scikit-Learn, OpenCV, FaceNet, Hugging Face Transformers
- Big Data & Analytics: Pandas, NumPy, Seaborn, Microsoft Excel
- APIs & Cloud Tools: OpenAI API, MIDI API
- Other Tools: Jupyter Notebook

Awards & Achievements:

- Harvard Hearsh System Innovation Lab Research Scholar/collaborator 2025
- Fulbright Award (2024–2026)
- 🏆 2nd Place Winner – MIT Media Lab Hackathon - Work Edition, May 2025

Work experience:

- Teaching Assistant-Computational Engineering| Worcester Polytechnic Institute | August 2025 – Current
-Support Instruction, help students understand C programming techniques. Facilitate Labs/Assignments: Provide hands-on support in lab sessions. Review and grade assignments or lab reports, office Hours.
- Project Analyst & Supervisor | TECCPROENG | New International Airport of Luanda | Oct 2022 – July 2024
- Managed and analyzed large-scale construction project data, improving operational decision-making.
- Designed scalable solutions for real-time project tracking and reporting.
- Service Desk Specialist | SINFIC SA, Luanda, Angola | July 2021 – Sept 2022
- Conducted computer hardware and software troubleshooting both remotely and in-person.
- Managed large GIS datasets, ensuring accurate mapping of polling stations for national elections.
- Applied data visualization and analytics using Excel to monitor electoral process efficiency.

Research & Projects:

1- AI-Driven Hypertension Detection from Voice Data (Thesis Research)

- Developed and evaluated ML models for hypertension detection using acoustic features.
- Designed ensemble MLP architectures with attention-based feature tokenization for improved calibration.

2- LoRA and QLoRA Fine-Tuning for Question Answering (CS 554 – NLP Coursework Project)

- Parameter-efficient fine-tuning of transformer-based LLMs using LoRA and Quantized LoRA for QA tasks.
- Compared scaling strategies (α/r) and rank-stabilization approaches to enhance training stability/consistency.
- Optimized tokenization, padding, and EOS handling to reduce loss variance and overfitting.
- Evaluated models with Exact Match (EM) and F1 scores, analyzing training curves and model generalization.

3- Retrieval-Augmented Generation for Biomedical Data (CS 554 – NLP Coursework Project)

- Built a hybrid dense + sparse retrieval pipeline combining FAISS (SBERT) and BM25 on the BioASQ dataset.
- Implemented RAG variants: *Query Transformation*, *HyDE (Hypothetical Document Expansion)*, *Self-RAG*, and *Semantic Chunking*.
- Integrated cross-encoder reranking, contextual compression, and adaptive retry policies for faithful evidence-based answers.
- Evaluated using BERTScore and LLM-as-a-judge comparisons, achieving improved accuracy and reduced hallucination rates.
- Fine-tuned and deployed Gemma-3-1B via LangChain and Hugging Face for scientific retrieval and QA tasks.

Languages:

Portuguese (Native) | English (Fluent) | Spanish (High Intermediate) | French (Intermediate)