

# JOSEPH MARTINEZ

+1 (757) 987-4351 | josephms957@gmail.com | Washington, DC | linkedin.com/in/josephmars | josephmars.me/

## PROFESSIONAL SUMMARY

---

Detail-oriented Data Scientist with 4 years of experience specializing in Machine Learning, LLM training and deployment, and simulation modeling. Proven ability to deliver actionable insights and improve processes through data-driven strategies and storytelling.

## PROFESSIONAL EXPERIENCE

---

### Inter-American Development Bank

Washington, DC, USA

*Generative AI Technical Lead Consultant*

*May 2025 - Present*

- Led the design and implementation of scalable Generative AI architectures using Azure, enabling efficient and secure integration of AI solutions within a complex technical environment.
- Provided technical leadership and mentorship across cross-functional teams, upskilling team members on best practices in Machine Learning, NLP, and responsible use of Generative AI.

### Aaru, Inc

New York, NY, USA

*AI Software Engineer*

*July 2024 - January 2025*

- Fine-tuned the state-of-the-art LLM Qwen 2.5 to deliver agentic, human-like responses, enhancing user satisfaction by 40% by leveraging a curated dataset of consumer preferences.
- Reduced GPU training and deployment costs by 60%, achieving \$3/hour, by setting up containerized LLM workflows with Runpod on custom hardware.
- Improved the accuracy of a generative agent-based model (gABM) by 30% through the implementation of live news and trends to the agents' memory and context window.

### Virginia Modeling Analysis and Simulation Center

Suffolk, VA, USA

*Data Scientist*

*September 2022 - December 2024*

- Trained Machine Learning classification models (RFC, ANN, and BERT) to identify types of frustration in more than 5000 news articles, achieving an average accuracy of 70%.
- Boosted GPT-3.5's ability to generate NetLogo procedural code by 50% by implementing Retrieval-Augmented Generation (RAG) techniques to support agent-based modeling.
- Improved domain-specific chatbot fidelity by 25% fine-tuning Llama2 7B on 330 news articles, ensuring high-fidelity responses.

### Old Dominion University

Norfolk, VA, USA

*Data Scientist*

*July 2021 - January 2022*

- Pre-processed 15.5M multilingual Tweets via Twitter's API and Python, reducing data noise by 40%, facilitating better trend analysis
- Geolocated, filtered, and estimated the tone of 7.3 million Tweets by fine-tuning 15 NLP models (spaCy, BERT, RoBERTa) and statistical analysis.
- Developed dashboards with Tableau and Google Data Studio, providing actionable migration insights for a \$1.4M Minerva research project involving teams from 4+ countries.

### Universidad del Norte

Barranquilla, Colombia

*Data Scientist*

*November 2020 - June 2021*

- Processed and filtered 4,680 geo-tagged Tweets using Python and Twitter's API, identifying migration patterns and trends.
- Conducted sentiment analysis to track xenophobic trends in social media, providing actionable insights for policymakers

## EDUCATION

---

### Old Dominion University

August 2022 - December 2024

*Master's, Modeling and Simulation Engineering*

*GPA: 3.7*

### Universidad del Norte

January 2017 - March 2022

*Bachelor's, Industrial Engineering*

*GPA: 3.6*

## SKILLS

---

**Skills:** FastAPI, Python, R, SQL, Pandas, Pytorch, ANN, PCA, NumPy, scikit-learn, Transformers, Hugging Face, OpenAI, Tableau, Power BI, Shiny, Git, GitHub, Docker, LangChain, Agile, CI/CD, Microsoft Azure, Terraform

**Languages:** German, Spanish, English