

# JOSEPH MARTINEZ

[in josephmars](#) | josephms957@gmail.com | (757) 987-4351 | [josephmars.github.io](#)

## SUMMARY

Research Assistant with **4 years of experience** in Machine Learning, and Natural Language Processing. Proficient in training **LLMs**, Machine Learning algorithms, and building Simulation models. Strong communication skills through professional engagements. Proven ability in programming and large-scale data analytics to derive **valuable insights**. Collaborative team member with a track record of **successful collaboration** on international research projects.

## WORK & RESEARCH EXPERIENCE

### Aaru, Inc

*AI Software Developer*

New York City, NY

July 2024 – Present

- Improved the performance of multi-agent LLM systems by designing and implementing state-of-the-art algorithms in complex Agent-Based Simulation environments.
- Designed simulation environments integrating survey and historical data to model political and business scenarios, enabling the generation of actionable insights and predictive analytics.
- Conducted innovative research in machine learning and Natural Language Processing while collaborating with interdisciplinary teams to develop innovative AI-driven solutions for real-world challenges.

### Virginia Modeling, Analysis, & Simulation Center (VMASC)

*Graduate Research Assistant*

Suffolk, VA

Sep 2022 – Present

- Trained the Large Language Model (LLM) GPT-3.5's to create ABM Simulation models, by improving its ability to generate NetLogo interface and procedural code elements through few-shot prompting.
- Developed a conversational Chatbot expert in a custom domain enhancing answer fidelity by 90% by fine-tuning the LLM Llama2 7B on a dataset of 330 news articles.
- Trained and fine-tuned Machine Learning classification models (RFC, ANN, and BERT) to identify types of frustration in textual data, achieving an average accuracy of 70%.
- Designed and maintained detailed documentation of models, algorithms, and reports on GitHub repositories.
- Presented poster at a research conference in Pittsburgh, and attended others in Arlington, VA, and Oxford, OH.

### Old Dominion University

*Social Media Data Analyst (Consultant)*

Norfolk, VA

Jul 2021 – Jan 2022

- Geolocated, filtered, and estimated the tone of **7.3 million Tweets** through fine-tuning fifteen NLP models (spaCy, BERT, RoBERTa) and Statistical Analysis. Models were highly reliable, with a mean accuracy of 72%.
- Extracted, cleaned, and pre-processed **15.5 million Tweets** in multiple languages with Twitter's API and Python.
- Collaborated with teams from Norway, Colombia, Greece, and USA on a \$1.4M USD Minerva research initiative.

### Universidad del Norte

*Social Media Data Analyst (Research Assistant)*

Barranquilla, Colombia

Nov 2020 – Jun 2021

- Trained a Natural Language Understanding (NLU) model with **IBM Watson** to semi-automate the identification of actors, factors, and relationships from news articles, achieving a final accuracy of 70%.
- Gathered, filtered, and cleaned geo-tagged 4,680 Tweets on migration using Twitter's API and Python.
- Conducted a lexicon-based Sentiment Analysis of Tweets to spot xenophobic trends.

## SKILLS

**NLP:** scikit-learn, Transformers, spaCy, Hugging Face, OpenAI GPT-4

**Machine Learning:** Classification, Regression, Clustering, PCA, NumPy

**Data Visualization:** Tableau, Power BI, R Shiny

**Deep Learning:** Pytorch, Keras, ANN

**Data Analysis:** Python, Pandas, R, SQL

**DevOps:** Git, GitHub, Docker

## EDUCATION

### Old Dominion University

M.Sc. Modeling and Simulation Engineering - GPA: 3.7

Norfolk, VA

Aug 2022 – Dec 2024

### Correlation One

Data Science for All (Bootcamp)

New York City, NY

Mar 2022 – Jul 2022

### Universidad del Norte

B.Sc. Industrial Engineering - GPA: 3.6

Barranquilla, Colombia

Jan 2017 – Mar 2022