

JOSEPH MARTINEZ

+1 (757) 987-4351 | josephms957@gmail.com | Norfolk, VA, USA | linkedin.com/in/josephmars/

PROFESSIONAL SUMMARY

Data Scientist with 4 years of expertise in Machine Learning and Natural Language Processing. Skilled in training and deploying LLMs and building simulation models. Focused on transforming data into impactful solutions, showcased through my work on high-profile research initiatives and contributions to interdisciplinary teams.

PROFESSIONAL EXPERIENCE

Aaru, Inc

New York, NY, USA

AI Software Developer

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 and Simulation Center

Suffolk, VA, USA

Graduate Research Assistant

September 2022 - December 2024

- 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

Norfolk, VA, USA

Social Media Data Analyst (Consultant)

July 2021 - January 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.
- Created dashboards with Tableau and Google Data Studio to represent migration data.
- Collaborated with teams from Norway, Colombia, Greece, and USA on a \$1.4M USD Minerva research initiative.

Universidad del Norte

Barranquilla, Colombia

Social Media Data Analyst (Research Assistant)

November 2020 - June 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.

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: Python, R, SQL, Pandas, Pytorch, ANN, PCA, NumPy, scikit-learn, Transformers, Hugging Face, OpenAI, FastAPI, Tableau, Power BI, Shiny, Git, GitHub, Docker

Languages: German, Spanish, English